

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:18:37 ; Search time 21 Seconds  
(without alignments)  
1543.645 Million cell updates/sec

Title: US-10-025-730-1  
Perfect score: 1704  
Sequence: 1 MKKMLFSGSKHNPAAIVKI.....FADEKNYLIKQIRDLKXTAP 337

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1376	80.8	341 2	I57997
2	1063.5	62.4	377 2	Ti6651
3	1006.5	59.1	338 2	T72129
4	834.5	49.0	329 2	T50117
5	685	40.2	305 2	G71441
6	632	37.1	348 2	B44448
7	485	28.5	399 2	S34681
8	143.5	8.4	339 2	T33477
9	134.5	7.9	677 2	H64574
10	128	7.5	430 2	H64709
11	125.5	7.4	298 2	B71695
12	125.5	7.4	1642 2	T08880
13	123.5	7.2	1285 2	B72420
14	120	7.0	1175 2	F64489
15	118.5	7.0	959 2	T00246
16	115	6.7	474 2	T43446
17	113.5	6.7	833 2	S71322
18	112.5	6.6	1411 2	S55123
19	111.5	6.5	725 1	J5016
20	111.5	6.5	2401 2	T28676
21	111	6.5	2166 2	G70163
22	111	6.5	2819 2	A90551
23	109.5	6.4	457 2	C82911
24	109.5	6.4	978 2	A70387
25	109.5	6.4	1830 2	E82909
26	109	6.4	695 2	T07283
27	109	6.4	1401 2	S11527
28	108.5	6.4	442 2	T18507
29	108.5	6.4	952 2	T50451

30	108.5	6.4	1163 2	D64315
31	108	6.3	568 2	S73254
32	107.5	6.3	483 2	I40055
33	107.5	6.3	855 2	E30106
34	107.5	6.3	1042 2	G64514
35	107.5	6.3	1726 1	SAZQGM
36	107.5	6.3	1726 2	A45948
37	107	6.3	570 2	S68686
38	107	6.3	1173 2	T43527
39	107	6.3	1727 2	T50073
40	106	6.2	474 2	S56748
41	106	6.2	1295 2	T24587
42	105.5	6.2	781 2	T00456
43	105.5	6.2	847 2	A56039
44	105.5	6.2	1091 2	T34107
45	105.5	6.2	1619 2	T18499

## ALIGNMENTS

### RESULT 1

I57997  
hypothetical calcium-binding protein - mouse  
C:Species: Mus sp. (mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 19-May-2000  
C:Accession: I57997  
R;Miyamoto, H.; Matsushiro, A.; Nozaki, M.  
Mol. Reprod. Dev. 34, 1-7, 1993  
A;Title: Molecular cloning of a novel mRNA sequence expressed in cleavage stage mouse  
A;Reference number: I57997; MUID:93119656; PMID:8418809  
A;Accession: I57997  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-341 <RES>  
A;Cross-references: GB:S51858; NID:G262933; PIDN:AB24801.1; PID:G262934  
C;Superfamily: Saccharomyces hypothetical protein YKL189W  
C;Keywords: calcium binding

Query Match 80.8%; Score 1376; DB 2; Length 341;  
Best Local Similarity 80.7%; Pred. No. 7.9e-95;  
Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;

Qy	4	MPL-FSKSHKNPAIVKILNDLAIKQ---	DKKTDKASEVSKSLOAMKELCGNEK	59
Db	1	MPPEFGSKSPADIVKLNKESNAVLEKQ	ISDKKAEKATEEVSKNLVAMKEILYGTNEK	60
Qy	60	EPPEVAQAQAEIYSSGLLVTLIADQL	IDFECKDVTQIFNNILARQIGTRSPTEYI	119
Db	61	EPQTEVAQAQAEIYNSGLLGLVADL	QLIDFECKDVAQIFNNILRRQIGTRPTVEYI	120
Qy	120	SAHPHILFMLLKGYEAPQIALRCGIM	LRCEIRHPLAKIILFNSQFRDFFKYVELSTFDI	179
Db	121	CTQQRILFMLKGYSEFIALNCGIM	LRCEIRHPLAKIILWSQFDFRYVEMSTFDI	180
Qy	180	ASDAFAFPKOLLTRHKVLVADFL	EQNYDTIPEDVEKLLQSNYVTKRQSLKGLGELIDR	239
Db	181	ASDAFAFPKOLLTRHKLSAEFL	EQHYDRFPSEYVEKLHSENYVTKRQSLKGLGELIDR	240
Qy	240	HNFAIMTKYISKPENLKMQLM	LLDKSNIOFEAFHVKFVFVAPSHKTPQIVEILLKNOP	299
Db	241	HNFTIMTKYISKPENLKMMLN	LLDKSNIOFEAFHVKFVFVAPNKTPQIDILLKNQT	300
Qy	300	KLIEFLSPQKERTDEQFADEKN	YLIKQIRDLAKTA	336
Db	301	KLIEFLSPQNDRTDEQFNDEK	YLVKQIRNLKRAA	337

### RESULT 2

Ti6651  
hypothetical protein R02E12.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000



C;Accession: T16651  
R;Leimbach, D.  
submitted to the EMBL Data Library, April 1996  
A;Description: The sequence of C. elegans cosmid R02E12.  
A;Reference number: Z18534  
A;Accession: T16651  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-377 <LEI>  
A;Cross-references: EMBL:U53337; NID:G1255833; PID:G1255838; PID:AAA96187.1; GSPDB:GN000020  
A;Experimental source: strain Bristol N2; clone R02E12  
C;Genetics:  
A;Gene: CESP:R02E12.2  
A;Map position: X  
A;Introns: 37/3; 146/2; 225/1; 315/3  
C;Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 62.4%; Score 1063.5; DB 2; Length 377;  
Best Local Similarity 60.5%; Pred. No. 7e-64;  
Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;

Qy 4 MP-LFSGSHKPAEIVKILKONLAILEK-----ODKTKDKASEEVSLSQAM 49  
Db 1 MPLLFGSKHSPADVVKTLREVITLTKLPPKLDKQGNIQSKYDKALDEVSKNAMI 60

Qy 50 KEILCGTNEKEPTE---AVAQAELYSGLLVTLADIQLIDFEGKQDVTOIFNNILR 106  
Db 61 KSPYIGNDSAPESSEHVQVQAQAQAEVYNANILPMLIKMLPKFEFECKDVGQIFNNLLR 120

Qy 107 RQGTGTRSPVVEYISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFR 166  
Db 121 RQGTGTRSPVVEYLGARPEILIQVQGVSDPIATCGMLRRESIRHDLAKIILYSDFY 180

Qy 167 DFFKYVLSLTFDIASDAFAATFKKLTTRHKLVLADFLQNYDTTFPEYKLLQSENYTKR 226  
Db 181 TFFLYVQSEVFDISSDAFSPFKELTTRHKAIIAEFLSDNYDTPFAQYONLNSKNVYTR 240

Qy 227 QSLKLGELLDNRNFNTMTYIKSPENKLMNLLRDKSPNIQFAHFVKFVSPHK 286  
Db 241 QSLKLGELLDNRNFNTMTYIKSPENKLMNLLRDKSPNIQFAHFVKFVSPHK 300

Qy 287 TQPIVEILLKNQPKLIFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 335  
Db 301 PKPISDILLNRNREKLVFELSFHNDRTDDEQFNDEKAYLIKQIRDLKKT 349

RESULT 3  
T27129  
hypothetical protein Y53C12A.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C;Accession: T27129  
R;Kershaw, J.; Lennard, N.  
submitted to the EMBL Data Library, September 1997  
A;Reference number: Z20315  
A;Accession: T27129  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-338 <WIL>  
A;Cross-references: EMBL:Z99277; PIDN:CAB16486.1; GSPDB:GN000020; CESP:Y53C12A.4  
A;Experimental source: clone Y53C12A  
C;Genetics:  
A;Gene: CESP:Y53C12A.4  
A;Map position: 2  
A;Introns: 29/3; 103/3; 136/2; 215/1; 282/3  
C;Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 59.1%; Score 1006.5; DB 2; Length 338;  
Best Local Similarity 57.2%; Pred. No. 4e-60;  
Matches 191; Conservative 60; Mismatches 78; Indels 5; Gaps 1;

Qy 5 PLFSGSHKPAEIVKILKONLAILEK-----ODKTKDKASEEVSLSQAMKEILCGTNEK 59  
Db 1 MPLLFGSKHSPADVVKTLREVITLTKLPPKLDKQGNIQSKYDKALDEVSKNAMI 60

Qy 50 KEILCGTNEKEPTE---AVAQAELYSGLLVTLADIQLIDFEGKQDVTOIFNNILR 106  
Db 61 KSPYIGNDSAPESSEHVQVQAQAQAEVYNANILPMLIKMLPKFEFECKDVGQIFNNLLR 120

Qy 107 RQGTGTRSPVVEYISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFR 166  
Db 121 RQGTGTRSPVVEYLGARPEILIQVQGVSDPIATCGMLRRESIRHDLAKIILYSDFY 180

Qy 167 DFFKYVLSLTFDIASDAFAATFKKLTTRHKLVLADFLQNYDTTFPEYKLLQSENYTKR 226  
Db 181 TFFLYVQSEVFDISSDAFSPFKELTTRHKAIIAEFLSDNYDTPFAQYONLNSKNVYTR 240

Qy 227 QSLKLGELLDNRNFNTMTYIKSPENKLMNLLRDKSPNIQFAHFVKFVSPHK 286  
Db 241 QSLKLGELLDNRNFNTMTYIKSPENKLMNLLRDKSPNIQFAHFVKFVSPHK 300

Qy 287 TQPIVEILLKNQPKLIFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 335  
Db 301 PKPISDILLNRNREKLVFELSFHNDRTDDEQFNDEKAYLIKQIRDLKKT 349

Db 4 PLFGKADKTPADVVKVKNLRDALLVIDRHGTNTSERKVEKAIBETAKWLAKALATFIYGSAN 63  
Qy 60 BPPTFAVLAQELYSGLLVTLADIQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEY 119  
Db 64 EPNQVQTQLAQEVYNANVLPMLIKHLKHFPECKDVASVFNLLRRQIGTRSPVVEY 123  
Qy 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYLVELSTFDI 179  
Db 124 ARPELITLLILGYEQPDIALTCGSMLEAREVHEHLARILVLYSEYFORFFVQSDVFDI 183  
Qy 180 ASDAFATPKDLTRHKLVLADFLQNYDTTFPEYKLLQSENYTKRQSLKLGELLDNR 239  
Db 184 ATDAFSTPKDLMTKHKMCAEYLDNDRFFQGYSAITNSENYVTRRQSLKLGELLDNR 243  
Qy 240 HNFATMTKYISKPENIKLMNLLRDKSPNIQFAHFVKFVSPHKTOPIVEILLKNOP 299  
Db 244 HNFSTMTKYITSPENIKLTVMLLRDRNRNIQFAHFVKFVSPHKTOPIVEILLKNOP 303  
Qy 300 KLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLK 333  
Db 304 KLIEFLTAFHNDRTDDEQFNDEKAYLIKQIRDLK 337

RESULT 4  
T50117  
mo25 homolog [imported] - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 28-Jul-2000  
C;Accession: T50117  
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, February 2000  
A;Reference number: Z25039  
A;Accession: T50117  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-329 <SEE>  
A;Cross-references: EMBL:AL157734; PIDN:CAB57774.1; GSPDB:GN000066; SPDB:SPAC1834.066  
A;Experimental source: strain 972h(-); cosmid c1834  
C;Genetics:  
A;Gene: SPDB:SPAC1834.066  
A;Map position: 1  
A;Introns: 34/3; 185/3  
C;Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 49.0%; Score 834.5; DB 2; Length 329;  
Best Local Similarity 51.5%; Pred. No. 1.2e-48;  
Matches 169; Conservative 63; Mismatches 93; Indels 3; Gaps 2;

Qy 6 LFSKSHKPAEIVKILKONLAILEK-----KODKTKDKASEEVSLSQAMKEILCGTNEKEPTE 64  
Db 4 LFNKRPKSTQDVVRCVCLDNLPKLEINNDKK--KSFEEVSKCLQNLVSLCGTAEEVPEAD 61

Qy 65 AVAQAELYSGLLVTLADIQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEYISAHPH 124  
Db 62 LVSDLSFQIYOSNLPFLVRLPKLEPEKSDKDTGLIFALLRRHVASRYPTVDYMLAHQ 121

Qy 125 ILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYLVELSTFDIASDAF 184  
Db 122 IFPVIVSYRYQEVAFVAFAGSILRECSRHEALNVLNSRDFWTFESLIQASSFDWASDAF 181

Qy 185 ATFKDLTRHKLVLADFLQNYDTTFPEYKLLQSENYTKRQSLKLGELLDNRNFAL 244  
Db 182 STFKSILLNHNKSKVAEFTSVHFDFEFKQYTVLLKSENYVTRRQSLKLGELLDNRNRSV 241

Qy 245 MTKYISKPENIKLMNLLRDKSPNIQFAHFVKFVSPHKTOPIVEILLKNOPKLEF 304  
Db 242 MTRYISSAENIKLMNLLRDKSPNIQFAHFVKFVSPHKTOPIVEILLKNOPKLEF 301

Qy 305 LSSFOKERTDDEQFADEKKNYLIKQIRDL 332  
Db 302 LSAFHTDRNDEQFNDEKAYLIKQIRDL 329



Query Match 28.5%; Score 485; DB 2; Length 399;  
Best Local Similarity 33.0%; Pred. No. 3.6e-25;  
Matches 113; Conservative 75; Mismatches 138; Indels 16; Gaps 6;

QY 7 FSKSHKNPAEIVKILKDNLAILEK---ODKTDKASBEVSKLQAMKEILGCTNEKEPP 62  
Db WKNKPTSPYAKLITIEQLNKSSPSLTQDNKR-KVQECTKVLITGTHKFIIVGDDTDPHT 74

QY 63 TEAVAQAQLYSSGLLVLTADLQIDFEKGKDVTFQFNILRRQIGTRSPTEVYISAH 122  
Db PEADIDELTAMRADVYFELLHFVLEFEARRECMILFISICLGYSKDNKFTVDVILVQ 134

QY 123 PHILFMLKGYE-----APQIALRCGIMLRECIHEPLAKILFNSQFRFFKVELS 175  
Db PKTISIMLRATAEVALQKGCQDIFLVGNMIECIKYEQLCRILKDPQLWFFEFKLG 194

QY 176 TPDIASDAPATFDKLLTRHKLVA-DPL--EQNYDTIFEDYEKLQSENVTYKRSQKL 232  
Db NFEISTESQILSAATAPKLVSKFFNEINIRFIKINKLMAHSGSVYKQSTKLL 254

QY 233 GELIDRHPAINTKYISKPENKLMNLLRDKSPNIQFEAFHVKVVFVASFPHKTOPIVE 292  
Db ASLIVRSNNALNIYINSPENKLTMTLMDTKSKNLQLEAFNVFKVMVAPRKRPFVD 314

QY 293 ILLKQNPKLIEFLSSQKERTDDEQFADEKNYLIKQIRDLKX 334  
Db ILVKNRDKLTYKTFGLD-SQDSTFDEREFIVQEIDSLPR 355

RESULT 8  
T33477  
hypothetical protein T27C10.3 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T33477  
R;Zhu, H.J.; Graves, T.; Hawkins, M.  
submitted to the EMBL Data Library, October 1998  
A;Description: The sequence of *C. elegans* cosmid T27C10.  
A;Reference number: Z21354  
A;Accession: T33477  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-339 <ZHU>  
A;Cross-references: EMBL:AF098504; PIDN:AAC67411.1; GSPDB:GN00019; CESP:T27C10.3  
A;Experimental source: strain Bristol N2; clone T27C10  
C;Genetics:  
A;Gene: CESP:T27C10.3  
A;Map position: 1  
A;Introns: 72/3; 120/3; 233/3; 295/1

Query Match 8.4%; Score 143.5; DB 2; Length 339;  
Best Local Similarity 19.3%; Pred. No. 0.02;  
Matches 38; Conservative 50; Mismatches 76; Indels 33; Gaps 4;

QY 159 ILFNSQFRDPFKYVELSTFDIASDAFAFKDLLTRHKLVDLFQNVDTIFEDYEKLQ 218  
Db LMNTNKFDP-----FDVIQGTFTDQIIFFTNHESANNFINNLPQMTLKLIA 150

QY 219 SENYVTRQSKLKLGLILDRHNPATMYTISKPENKLMNLLRDKSPNIQFEAFHVK 278  
Db CSNFFIQAKSKFKNELFTAGTNYETRSLMMAEPAFIKLVLAQSNKHAVSRVAVSILE 210

QY 279 VFVASFPHKTOPIVELLLKNQKLLIEFL-----SSQKERTDDEQAD----- 320  
Db IFIRNPRNSPEVHEFIGRNRNVLIAPFNFSAPIHYQGSNPEKE----DAQVARMAYKLLN 267

QY 321 ---EKNYLIKQIRDLKX 334  
Db WDMQRPFTEQQLQDFEE 284

RESULT 9  
H64574

DNA topoisomerase I - *Helicobacter pylori* (strain 26695)  
C;Species: *Helicobacter pylori*  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C;Accession: H64574  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C  
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: H64574  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-677 <TOM>  
A;Cross-references: GB:AE000559; GB:AE000511; NID:92313536; PIDN:AAD07502.1; PID:923135  
C;Superfamily: DNA topoisomerase I

Query Match 7.9%; Score 134.5; DB 2; Length 677;  
Best Local Similarity 21.6%; Pred. No. 0.19;  
Matches 88; Conservative 58; Mismatches 134; Indels 127; Gaps 16;

QY 7 FSKSHKNPA-EIVKILKONL-----AILEKQDKK---TKASEVSKSLQAMKE 51  
Db 222 FXFKDKNEASQFLKDKLKDGLGSMVLSKESLSNKKPKPKFTTSKLLSQASKSLKI--- 278

QY 52 ILCGTNEKEPPTPEAVAQIAQLYSSGLLVLTADLQIDFEKGKDVTFQFNILRRQIGT 111  
Db 279 -----PTKRIQAQLAKLFEAGLITYHTTDSFELSPEYLKEHEVFEFIY----- 322

QY 112 RGPTV-----EXIS-----AHPHILFMLKGYEAPQIALRCGIMLRECIH 153  
Db 323 --PSVQVQREYKAGKNSQAEAEHAIIRTHPALKOLEKVCSDAKISEELAKLYLIYTN 380

QY 154 PL---AKIILFNSQFRDPFKYVELSTFDIASDAFAFKDLLTRHKLVDLFQNVDTIF 210  
Db 381 TICSQSRNALY-NOYDCIEK-----IKSESFLSPKLLKEKGFLEELIQGKEEIN 431

QY 211 EDYEKLQSENVTYKQSKLKLGLILDRHNPATMYTISKPENKLMNLLRDKSPNIQ 270  
Db 432 RE--EOSEIENFSLKENDSVPLKEVFIKK-----IEKSPKPKYESAFPLDLESEG--- 481

QY 271 FEAFHVKVVFVASFPHKTOPIVEILLKNQ-----PKLIEFLSSQKERTD--- 315  
Db 482 -----IGRPSYASFLDLLKRYISIDTKNATPTTSQGLEVISFFPKDKVEVD 531

QY 316 -----EQF-----ADEKNYLIKQIRDLKTA 336  
Db 532 IALTSKDKSKLGNNTKQFEELDLIMRGEASVEKFMLEVISKLKSTA 578

RESULT 10  
H64709  
hypothetical protein HP1520 - *Helicobacter pylori* (strain 26695)  
C;Species: *Helicobacter pylori*  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C;Accession: H64709  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C  
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: H64709  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-430 <TOM>  
A;Cross-references: GB:AE000650; GB:AE000511; NID:92314700; PIDN:AAD08565.1; PID:923147  
C;Superfamily: *Helicobacter pylori* hypothetical protein HP1520

Query Match 7.5%; Score 128; DB 2; Length 430;  
Best Local Similarity 20.9%; Pred. No. 0.29;

Matches	82;	Conservative	73;	Mismatches	135;	Indels	102;	Gaps	20;	
QY	7	FSKSHKNPAE	----	VKILKDNLAILEK	OKDKTDKASEE	VSKSLQAMKEIL	CGTNEKEPP	62		
Db	60	FYPNRKSKIE	IEFNGE	KILKENVAVF	HSYDE--	EFSSDSVTTFMA	KSDL-----	KQY 111		
QY	63	TEAVAQAE	LYSGGLVTU	--TA-----	DIQLDFEG	KDQVTFQFNILR	-----	106		
Db	112	DNILLEKE	-KKALLKSL	RDIASGPDYEE	EIKTKIKNEK	KSFYELDNHLE	TESSEK	169		
QY	107	-----	ROI	GTSPFTV-EYIS	APHILFMLL	KGYEARQIAL	RCGIMLRCIR	HEPLAKII 159		
Db	170	HYSFKYRD	IFDGSKVK	QDFYNGHDL	EYFNKYQ-----	-----	-----	ELLSQSK 211		
QY	160	LF-----	SNQ	RDFEFK	VELSTFDIAS	DAFATFKOLL	TRHKVLVADFLEQ-----	204		
Db	212	IFKHNSG	DGTWAD	LKALENNR	FFKANHS	LIAGEELTN	VQKL-SDIFENE	KNRIL 270		
QY	205	NYDTIFED	YKELQSEN	VYTKROSL	KLGL-----	-----	LDRHNF--	AIMTKYISKP 252		
Db	271	NNEELKES	FPKI--	EKVINANK	ELKAFKDA	ISKNDNTLL	TEFLDYDSFRK	KVLFSYLKQV 327		
QY	253	ENLKM	MLLRDK	SPNIQFEA	HVFKV	FVASPHK	QTPIVEILL	KNQPKLIEFL	SSQKE 311	
Db	328	IQNVSL	NVLYRE	KKCP	IE-----	EILKQASK	QDKQEWES	VEIEF--	NORFLVPFKVELQNQ 381	
QY	312	R-----	TDDEQ	----	FADEKN	YLKIQIRDL	KK 334			
Db	382	KOILLNK	DAQFR	IFSD	NDQDMNV	QKEDIQK 413				
RESULT 11										
B71685										
hypothetical protein RP295 - Rickettsia prowazekii										
C:Species: Rickettsia prowazekii										
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000										
C:Accession: B71685										
R:Anderson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; A										
Nature 396, 133-140, 1998										
A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitoch										
A:Reference number: A71630; MUID:99039499; PMID:9823893										
A:Accession: B71685										
A>Status: preliminary; nucleic acid sequence not shown; translation not shown										
A:Molecule type: DNA										
A:Residues: 1-298 <AND>										
A:Cross-references: GB:A7235271; GB:A7235269; NID:G3868717; PIDN:CAA14756.1; PT										
A:Experimental source: strain Madrid E										
C:Genetics:										
A:Gene: RP295										

Query Match	7.4%	Score 125.5	DB 2	Length 298
Best Local Similarity	20.1%	Pred. No. 0.27		
Matches	62	Conservative 57	Mismatches 114	Indels 75
Gaps	13			
Qy	73	LYSGSLVTLTADIQLIDFEKKDVQ	-----	IFNNILRRIGTFS 113
Db	6	LFIQLLIVTSLVKAEIIEVDSLINKITQDFKVNYNKYLPODLLVVTLDKFLFSFGV	-----	63
Qy	114	PTVEVISASHPHILFWLLKGY--EAQIARLCGIMLRCEIRHEPLAKIILFSNCP	-----	166
Db	64	PIGEYIDQRYIALAPLFSHINKPKIY	-----	109
Qy	167	-DFFKYV-ELSTFDI-----ASDAFAFKDLTRHKVLVADFLQNVDTIFEDYEKLLQSE	-----	220
Db	110	SDFNFVNEMNSQIPITAVNNGFTGNFNIPKFLWFADYLLKNF---	YIDFSKSPNN 166	
Qy	221	NYVTKRQSLKLGEILIDRHNFAIMTKYISKPNL---	KLMNNILRDKSPNTQFAFHV 277	
Db	167	NYI-----IFNNLDSFNTVPYFKGLTSTNNIPASKVILNF---	IQINFIPCK 213	
Qy	278	KVFVASPHKTPQIVELLKNQPKLIEFSSF--QXERTDDEQFADEKNY	-----	LIKOI 329
Db	214	FTLITSSRELLRSMFQNLNYSNLIIFYGHYNNKSI	ISDDKDYKQIAYYTKRMINDLPI 273	

QY 330 RDLKKTAP 337  
DB 274 NKLKRNPP 281

RESULT 12

TM0880  
C:Species: Homo sapiens (man)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T08880  
R:Lin, J.W.; Wysocki, M.; Madhavan, R.; Sealock, R.; Kim, J.U.; Sheng, M.  
J. Neurosci. 18, 2017-2027, 1998  
A:Title: Yotiao, a novel protein of neuromuscular junction and brain that interacts with  
A:Reference number: Z16511; MUID:98151389; PMID:9482789  
A:Accession: T08880  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1642 <LIN>  
A:Cross-references: EMBL:AF026245; NID:g2623067; PIDN:AAB86384.1; PID:g2623068  
C:Genetics:  
A:Map position: 7q21-22  
C:Keywords: brain; cerebral cortex; coiled coil; neuromuscular junction; skeletal muscle

Query Match 7.4%; Score 125.5; DB 2; Length 1642;  
Best Local Similarity 20.2%; Pred. No. 2.4; Indels 115; Gaps 15;  
Matches 77; Conservative 73; Mismatches 117

QY 18 VKIKONLAILEKQDKKTKASEEVSKSQAMKEILCGTNEKEPPTTEAVAQLAQELYSYG 77  
DB 664 IEKLNKNGIHYKQ-QIDGLQNMESQKIETWQ-----FEKDNLTQKQQLILE----- 710

QY 78 LVTILADLQ-LIDPEKQDVTFINNIIRROI-----GTSPVTEVISAHPI 125  
DB 711 --ISKLDLQOOSLVNSKSEMTLQI--NELQKEIILROEKEKGTLEQVQLKTEL 766

QY 126 LFMLLKGYEAPQIALRCGIMLCIRHEPLAKILFNSQFRDFKVKYVELSTFDIASDAFA 185  
DB 767 LEQMKKE-----NDLQKFAQLEAEN-SILKDEK 797

QY 186 TFKDLALTRH-----KVLVADFLQ-QNYTFIDYKLLQSNVYTKRSLKLGELIL 237  
DB 798 TLEMLKTHTPVQSDEERLIFLDSIKSKSKDSVWEKETEILIEENEDLKQOQIQLNESIEK 857

QY 238 DRHNFALNTK-----YISKPENLKLMMNLIRD 264  
DB 858 QRNTFSFAEKNFVNYQELQEVYACLVKVDLSDSKNQELEYKSKKALNESHQRI 917

QY 265 KSPNIQFEA--FHVKVFVASPHKTPQIVELLKNQPKLIEFLSSFOKERTD-DEQFAD- 320  
DB 918 NPTIVKMKSSVDFDKTFVA--ETLEMGVEXKQDTTELMKLELVTKRKLELSQRLSDL 974

QY 321 -----EKNYLIKQIRDLK 333  
DB 975 SEQLKPKPGISFLNEEVKSLK 996

RESULT 13

B72420  
C:Species: Thermotoga maritima (strain MS88)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: B72420  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickie,  
Garrett, M.M.; Steward, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: B72420  
A:Status: preliminary  
A:Molecule type: DNA

A;Residues: 1-1285 <ARN>  
A;Cross-references: GB:AB001695; GB:AB000512; NID:G4980569; PIDN:AA035182.1; PID:G498057  
A;Experimental source: strain MGB8  
C;Genetics:  
A;Gene: TM0088

Query Match  
Best Local Similarity 7.2%; Score 123.5; DB 2; Length 1285;  
Matches 86; Conservative 78; Mismatches 129; Indels 107; Gaps 23;

QY 1 MKKMPFKSKHNPAEIVKILKDNLAILEKQD-----KKT-----DKASEV-----SKS 45  
DB 556 LKVAMLSGKEEN-----VQKAAEELQISSERIIRFVKKTENVIDKAKNVVLQLSVS 611  
QY 46 LQAMKEILCGTNEKEPPTAVLAQELYSGL-----LVTLIAD-- 85  
DB 612 IEEGLNELVIGERE-EVEKAAADLLQKFSSEVEISRDFVKLPWSWIDQEKLELVYKNSA 670  
QY 86 ---LQID---FECKD---VTOIENNLARQIG--TRSPTEVI---SAHPHILEWL 129  
DB 671 GITYEILDGVVFTGKENVKAKELSDIVEK-LGEVKETVTFLEWNSFPVDFEIN 729  
QY 130 LKGYEAPQIALRCGIMLRCEIRHEPLAKIIL-----FSNQPRDFP-----KYVELST 176  
DB 730 LSGKLYPDVT-----CFSLDQLGLLVKGSSEAVEDELSSMYRSFPERHOKIVKENV 780  
QY 177 PD---IASDAPATFKDOLLTRHKVLVADFLQNYDTIFEDYEKLQSENVTYKRSKLKLG 233  
DB 781 PDRLMLEVPSGFSFEFTFLVLVPEVKQ-----VYLDKMLNLVVEVPVQSOSERVKSL 836  
QY 234 EILDRHNFAITWKYIS-----KPNEL-KLMNLLRDKSPNTOFEAF-HVFVKVVAS 283  
DB 837 DFELKEEAVSEKAVKSVTIPSGVNPDELSSYLKLLR-----NVEITVFPNMGQIVEG 892  
QY 284 P-HKTQPIVEILLKNQPKLIEFLSSFKQKERTDDEQFADEK 322  
DB 893 PENEVEKAVELVEAEKIV-----LKERKDYVKVSDGK 926

RESULT 14  
F64489  
Hypothetical protein MUI519 - Methanococcus jannaschii  
C;Species: Methanococcus jannaschii  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C;Accession: F64489  
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Reson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996  
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R. Science 273, 1058-1073, 1996  
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A;Reference number: A64300; MUID:96337999; PMID:8688087  
A;Accession: F64489  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1175 <BUL>  
A;Cross-references: GB:U67593; GB:L77117; NID:G2826427; PIDN:AAB99538.1; PID:G1500409; T  
C;Genetics:  
A;Map position: FOR1494096-1497623

Query Match  
Best Local Similarity 7.0%; Score 120; DB 2; Length 1175;  
Matches 76; Conservative 58; Mismatches 131; Indels 88; Gaps 15;

QY 7 FSKSHKNPAEIVKILKD-NLAILEKQDKTDTKASEVSKSQAMKEILCGTNEKEPPTA 65  
DB 232 FNKFRFNQDFDKYLTDENIAFRPHVMKFDFAENIKKVIABLE-----GSKYKPLPG 287  
QY 66 VLAQELYSGLLVTLADLQIDFEGKKDVTQIENNLRRQIGURSPTEVISAHPHI 125  
DB 288 V-----LYFLGMEADYSRVIELWKNKEGEEKENYALJI-ESENKENLEF----- 333  
QY 126 LFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFSNQPRDFPK-----YVELSTFIDA- 180

DB 334 -----GITKKVIDKFAQKEEFREFLKNVAVYYELSAFKLEK 370  
QY 181 -----SDAFATFKDOLLTRHKVLVADFLQNYDTIFEDYEKLQSENVTYKRSQSL----- 229  
DB 371 IKEQVEKEFINLDNIKNPYILVED-LKEN-----DSFERIIFELDSWERRLLGDKENP 424  
QY 230 -----KLIGELILDRH-----NFAIMTKYISKPENLKLMMNLLRDKSPNTOFEAF----- 274  
DB 425 YSPYRVALLVE-ILKRLHSSGNTTISTK-----DLKOFFEKMDKQIVKITFEFLRII 477  
QY 275 HVFKVVFASPHKTOPIVEILLKNQPKLIEFLSSFKQKERTDDEQFADEKNYLIK 327  
DB 478 BEYKDIIS--EKVEIVKKEVKNENKEIIEFTLKEIREVEEIIENTINYLLK 528

RESULT 15  
T00246  
DNA polymerase V - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 31-Jan-2000  
C;Accession: T00246; T39442  
R;Shimizu, K.  
A;Submitted to the EMBL Data Library, March 1998  
A;Description: S.pombe homolog of S.cerevisiae DNA polymerase V.  
A;Reference number: Z14129  
A;Accession: T00246  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-959 <SHI>  
A;Cross-references: EMBL:AB012696; NID:d1224325; PIDN:BAA32046.1; PID:d1033008  
R;Xiang, Z.; Aves, S.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
A;Submitted to the EMBL Data Library, March 1998  
A;Reference number: Z21854  
A;Accession: T39442  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-959 <LYN>  
A;Cross-references: EMBL:AL022305; PIDN:CAA18436.1; GSPDB:GN00067; SPDB:SPBC14C8.14c  
A;Experimental source: strain 972h; cosmid c14C8  
C;Genetics:  
A;Gene: pol5+; SPBC14C8.14c  
A;Map position: 2  
A;Introns: 56/3

Query Match  
Best Local Similarity 7.0%; Score 118.5; DB 2; Length 959;  
Matches 80; Conservative 63; Mismatches 135; Indels 113; Gaps 19;

QY 9 KSHKN-----PAEIVKILKONLAILEKQDKKTKASEVSKSQAMKEILCGTNEKEPP 62  
DB 522 KSPKNLLISMDSEVIEIVQKSLSVLHKVTKKIDKQAHL-QQLNAP----- 567  
QY 63 TEAVAQELYSGLLVTLADLQIDFEGKKD-VTQIFNNILRRQIGURSPTEVYI- 119  
DB 568 -----QLLYSVLLQVYAGDTSDIVLEDIDNCYSKVFNKSKRESTSNEPTAMEIL 619  
QY 120 -----SAHPHLF-----MLKCY-----EAPQIALRC-----GIMLRCEI 150  
DB 620 TEVMSLSLRPSLLRLKLVDMLTFTSFSDMNRSEIHLCDVLKAKESVKDSGMPAGEV 679  
QY 151 RHEPLAKIILFSNQPRDFKVKVVELSTFDIASDAFATFKDOLLTRHKVLVADFLQNYDTIF 210  
DB 680 EEDAFGE-----TEMDEDDFEEDTDEIEEQSD-----WEMISNQDASNEELERKLDKVL 730  
QY 211 ERYEKLQ-----SNVYTKRSQKL-----LGBLILDRHNFAIMTKYISKPENLKLMMNLL 262  
DB 731 EDADAKVKDESESEBELMNDQMLADEKLAIEVFRER-----KKASNKEKKKNAQ 780  
QY 263 RUKSPNIQEAHFVKV--FVASPHKTO-----PIVEILLKNQPKLIE----- 303  
DB 781 ETQKQIVQPKV-----KVIDLIDNYTKTPNNGLGFEFLPLELLEMLIKTKHKVLEKQAV 836

Search completed: April 12, 2004, 10:27:41  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 09:49:37 ; Search time 18 Seconds  
(without alignments)  
974.869 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 1704

Sequence: 1 MKKMPFLFSKHNPAIVKI.....FADEKNYLIKQIRDLKKTAP 337

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1685	98.9	334	1	MO2L_HUMAN	Q9h984 homo sapien
2	1669	97.9	334	1	MO2L_MOUSE	Q9db16 mus musculu
3	1381	81.0	341	1	MO25_HUMAN	Q9y376 homo sapien
4	1376	80.8	341	1	MO25_MOUSE	Q96138 mus musculu
5	1111	65.2	339	1	MO25_DROME	P91891 drosophila
6	1006.5	59.1	338	1	MO2M_CABEL	O18211 caenorhabdi
7	834.5	49.0	329	1	YFV6_SCHPO	Q9p7q8 schizosacch
8	776	45.5	321	1	DE76_CHLPR	Q9xfy6 chlorella p
9	728	42.7	343	1	MO2N_ARATH	Q9fgk3 arabidopsis
10	716.5	42.0	343	1	MO2M_ARATH	Q9m04 arabidopsis
11	666	39.1	384	1	HYMA_EMENI	O60032 emericelella
12	632	37.1	348	1	MO2L_ARATH	Q9zq77 arabidopsis
13	485	28.5	339	1	HYM1_YEAST	P32464 saccharomyc
14	143.5	8.4	339	1	MO2L_CABEL	Q9tzm2 caenorhabdi
15	128.5	7.5	391	1	AKA9_HUMAN	Q99996 h a-kinase
16	125.5	7.4	298	1	Y295_RICPR	Q9zdm2 rickettsia
17	120	7.0	1175	1	YF19_METJA	Q58914 methanococc
18	118.5	7.0	959	1	DPO5_SCHPO	O60094 schizosacch
19	116.5	6.8	724	1	HMNR_HUMAN	O75330 homo sapien
20	115	6.7	474	1	GSBH_MOUSE	P51855 mus musculu
21	112.5	6.6	1411	1	YM42_YEAST	O01214 saccharomyc
22	109.5	6.4	978	1	FA50_AQUAE	O67124 aquifex ao
23	109	6.4	695	1	YCX7_CHLVU	O20159 chlorella v
24	109	6.4	1401	1	LATA_LATWA	P23631 latrodectus
25	108.5	6.4	586	1	2A5D_HUMAN	Q28653 o serine/th
26	108.5	6.4	602	1	2A5D_HUMAN	Q14738 h serine/th
27	108.5	6.4	1075	1	Y124_METJA	O57588 methanococc
28	108	6.3	568	1	DNAB_PORPU	P51333 porphyra pu
29	107.5	6.3	483	1	ACPA_BACAN	Q44643 bacillus an
30	107.5	6.3	1042	1	TSRH_METJA	Q60295 methanococc
31	107.5	6.3	1726	1	MSP1_PLAFC	P04934 plasmodium
32	107.5	6.3	1726	1	MSP1_PLAFC	P50495 plasmodium
33	107	6.3	1727	1	ALM1_SCHPO	Q9utk5 schizosacch

```

34 106 6.2 474 1 GSHB_HUMAN P48637 homo sapien
35 105.5 6.2 793 1 REGA_DICDI Q23917 dictyostell
36 105.5 6.2 847 1 RSG2_RAT Q63713 rattus norv
37 105 6.2 767 1 GOA1_HUMAN Q92805 homo sapien
38 104.5 6.1 634 1 YG74_METJA Q59068 methanococc
39 104.5 6.1 1701 1 MSP1_PLAFC P13819 plasmodium
40 104.5 6.1 1701 1 MSP1_PLAFC P08869 plasmodium
41 104 6.1 859 1 MUTS_AQUAE O66852 aquifex ao
42 104 6.1 1290 1 RASO_SCHPO Q9utj8 schizosacch
43 104 6.1 1682 1 MSP1_PLAFC P19598 plasmodium
44 103.5 6.1 641 1 PRIM_UREPA Q9p26 ureaplasma
45 103 6.0 2663 1 CENE_HUMAN Q02224 homo sapien

```

#### ALIGNMENTS

```

RESULT 1
MO2L_HUMAN STANDARD; PRT; 334 AA.
AC Q9H984; Q9BZ33;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE MO25-like protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 4-334 FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE OF 276-334 FROM N.A.
RA Pearce A.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK022639; BAB14147.1; ALT_INIT.
DR EMBL; AL138875; CAC28084.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 334 AA; 38728 MW; 97702273D8548432 CRC64;

```

Query Match 98.9%; Score 1685; DB 1; Length 334;  
Best Local Similarity 99.7%; Pred. No. 1.3e-100;  
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 4 MPLFSKHNPAIVKILKONLAILEKQDKTKASEEYVSKSLQAMKEILCGTNEKEPPT 63
Db 1 MPLFSKHNPAIVKILKONLAILEKQDKTKASEEYVSKSLQAMKEILCGTNEKEPPT 60
QY 64 EAVAQAQLYSSGLLVTLIADQLIDFGKKDVDTQIFNNILRRQIGTRSPVTEYISAHP 123
Db 61 EAVAQAQLYSSGLLVTLIADQLIDFGKKDVDTQIFNNILRRQIGTRSPVTEYISAHP 120

```



QY 124 HILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQNRDFPKYVELSTFDIASDA 183  
 Db 121 HILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQNRDFPKYVELSTFDIASDA 180  
 QY 184 FATEKDLLTRHKVVLVADFLFQNYDTTFEDYEKLLQSENVTYKRSKILGELILDRHFA 243  
 Db 181 FATEKDLLTRHKVVLVADFLFQNYDTTFEDYEKLLQSENVTYKRSKILGELILDRHFA 240  
 QY 244 INTKYSKPNENKLMNLRDSDSPNQFEAFHVKFVSPHKTQPIVILLKNQPKLIE 303  
 Db 241 INTKYSKPNENKLMNLRDSDSPNQFEAFHVKFVSPHKTQPIVILLKNQPKLIE 300  
 QY 304 FLSSFOKERTDEQFADEKNYLIKQIRDLKKTAP 337  
 Db 301 FLSSFOKERTDEQFADEKNYLIKQIRDLKKTAP 334  
 RESULT 2  
 MO2L\_MOUSE  
 ID MO2L\_MOUSE STANDARD; PRT; 334 AA.  
 AC Q9DB16; Q8BG52; Q91WB8; Q91YL0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE MO25-like protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC STRAIN=C57BL/6J;  
 RC TISSUE=Cerebellum, Eye, Pituitary, and Testis;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schrim L.M., Kanapin A., Matsuda H., Batalov S., Beisler K.W.,  
 RA Blake J.A., Bradt D., Brusik V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustinchin S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=FVB/N; TISSUE=Mammary gland, and Salivary gland;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9DB16-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9DB16-2; Sequence=VSP\_007417, VSP\_007418;  
 CC Note=No experimental confirmation available;  
 CC -!- SIMILARITY: Belongs to the Mo25 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AK005323; BAB23953.2; ALT INIT.  
 DR EMBL; AK030474; BAC26978.1; ALT INIT.  
 DR EMBL; AK053642; BAC35457.1; ALT INIT.  
 DR EMBL; AK076758; BAC36470.1; ALT INIT.  
 DR EMBL; AK076867; BAC36513.1;  
 DR EMBL; BC016128; AAH16128.1;  
 DR EMBL; BC016546; AAH16546.1;  
 DR MGI; MGI:1916259; 1500031K13Rik.  
 DR InterPro; IPR008938; ARM.  
 DR Pfam; PF01204; Mo25; 1.  
 DR Alternative splicing.  
 DR VARSPLIC 276 293  
 VFWASPHKTQPIVEILLK -> NSVFITNRHGLKRWLSS  
 (in isoform 2).  
 /FTid=VSP\_007417.  
 Missing (in isoform 2).  
 S -> P (IN REF. 1; BAB23953).  
 L -> R (IN REF. 2; AAH16546).  
 QUERY MATCH 97.9%; Score 1669; DB 1; Length 334;  
 Best Local Similarity 98.5%; Pred. No. 1.4e-99;  
 Matches 329; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 MPLFSKSHKNPAIEVKILKNDLALAEKQDKTKDASEEVSXSLOAMKEILCGTNEKEPPT 63  
 Db 1 MPLFSKSHKNPAIEVKILKNDLALAEKQDKTKDASEEVSXSLOAMKEILCGTNDKEPPT 60  
 QY 64 EAVAQAQELYSGLLVTLIADLIQDFEGKQDVTOIFNNILRRQIGTSPVVEYISAPH 123  
 Db 61 EAVAQAQELYSGLLVTLIADLIQDFEGKQDVTOIFNNILRRQIGTSPVVEYISHP 120  
 QY 124 HILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQNRDFPKYVELSTFDIASDA 183  
 Db 121 HILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQNRDFPKYVELSTFDIASDA 180  
 QY 184 FATEKDLLTRHKVVLVADFLFQNYDTTFEDYEKLLQSENVTYKRSKILGELILDRHFA 243  
 Db 181 FATEKDLLTRHKVVLVADFLFQNYDTTFEDYEKLLQSENVTYKRSKILGELILDRHFT 240  
 QY 244 INTKYSKPNENKLMNLRDSDSPNQFEAFHVKFVSPHKTQPIVEILLKNQPKLIE 303

```

Db 241 IMTKYISKPNLKLWNLADKSPNIOFEAFHFVKFVASFPHKTOPIVEILLKNQPKLIE 300
QY 304 FLSSFQKERTDDEQFADEKKNYLKQIRDLKKTAP 337
Db 301 FLSSFQKERTDDEQFADEKKNYLKQIRDLKKAAP 334

RESULT 3
MO25_HUMAN
ID MO25_HUMAN STANDARD; PRT; 341 AA.
AC Q9Y376;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MO25 protein (CGI-66).
GN MO25.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
[2]
SEQUENCE FROM N.A.
RX TISSUE=Hypothalamus;
RA Jin W., Shi J., Ren S., Gu J., Fu S., Huang Q., Dong H., Yu Y., Fu G.,
RA Wang Y., Chen Z., Han Z.;
RT "A novel gene expressed in the human hypothalamus.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RX TISSUE=Ductenium;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohnyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
RA Schnorch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF151824; AAD34061.1; -
CC ENBL; AF153536; AAF14873.1; -
CC ENBL; BC020570; AAB20570.1; -
CC InterPro; IPR008938; ARM.

```

```

DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39869 MW; EC710A528B6F9811 CRC64;

Query Match 81.0%; Score 1381; DB 1; Length 341;
Best Local Similarity 81.0%; Pred. No. 3.2e-81;
Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;

QY 4 MPL-PSKSHKNPABIYKILKDNLAILEKQ---DKTKDASBEYSKGLQAMKEILCGTNEK 59
Db 1 MPFPFCKSHKSPADIYKLNKESNAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
QY 60 EPTEAVALAQLYSGLLVLIADLQIDPECKDVTOIFNNILRQIGTSPTVEYI 119
Db 61 EPQTEAVALAQLYNSGLLSTLVADLQIDPECKDVQIFNNILRQIGTSPTVEYI 120
QY 120 SAHPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFRDFPKYVELSTFDI 179
Db 121 CTQONILFLLKGYESPALNCGIMLRCEIRHEPLAKILFNSQFRDFPKYVELSTFDI 180
QY 180 ASDAFATFKDLLTRHKLVVADLEQNYDTIFEDYEKLLQSENYYTKRQSLKLGELILDR 239
Db 181 ASDAFATFKDLLTRHKLVVADLEQNYDTIFEDYEKLLQSENYYTKRQSLKLGELILDR 240
QY 240 HNFATMTKYSKPNLKLWNLADKSPNIOFEAFHFVKFVASFPHKTOPIVEILLKNQ 299
Db 241 HNFATMTKYSKPNLKLWNLADKSPNIOFEAFHFVKFVASFPHKTOPIVEILLKNQ 300
QY 300 KLIEFLSSFKERTDDEQFADEKKNYLKQIRDLKKTAA 336
Db 301 KLIEFLSSFKERTDDEQFADEKKNYLKQIRDLKKTAA 337

RESULT 4
MO25_MOUSE
ID MO25_MOUSE STANDARD; PRT; 341 AA.
AC Q06138;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MO25 protein.
GN MO25 OR CAB39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93119656; PubMed=8418809;
RA Miyamoto H., Matsushiro A., Nozaki M.;
RT "Molecular cloning of a novel mRNA sequence expressed in cleavage
RL stage mouse embryos.";
RL Mol. Reprod. Dev. 34:1-7(1993).
CC -!- FUNCTION: ONE OF THE FIRST GENES TO BE TRANSCRIBED DURING MOUSE
CC DEVELOPMENT. MAY PLAY SOME GENERAL FUNCTION.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- DEVELOPMENTAL STAGE: TRANSCRIBED DURING EARLY MOUSE DEVELOPMENT.
CC DETECTED AT ALL DEVELOPMENTAL STAGES FROM THE EGG THROUGH THE
CC BLASTOCYST, MOST ABUNDANT AT THE 2-CELL STAGE.
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S51858; AAB24801.1; -
CC MGD; MGI:107438; Cab39.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR004892; Mo25.

```

```

DR Pfam: PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39842 MW; E7F668529D6F8E911 CRC64;

Query Match
Best Local Similarity 80.8%; Score 1376; DB 1; Length 341;
Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;

QY 4 MFLFGKSHKQNPAAIVKILKNDLAILEKQ---DKKTDKASEEVSKSQAMKEILCGTNEK 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MPFPFGKSHKSPADIVKILKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
QY 60 EPPTREAVAQALQELYSGLLVTLIADQLIDFEGKKQVTFQIFNNILRRQIGTRSPTEVI 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 EPQTEAVAQALQELYSGLLVTLIADQLIDFEGKKQVTFQIFNNILRRQIGTRSPTEVI 120
QY 120 SAHPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNGQFRDFPKYVELSTFDI 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 CTQONILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNGQFRDFPKYVELSTFDI 180
QY 180 ASDAFATFQDLTRHKLKLVADFLQYVDFYEDYEKLQSENVTYTKRQSLKLLGELLDR 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 ASDAFATFQDLTRHKLKLVADFLQYVDFYEDYEKLQSENVTYTKRQSLKLLGELLDR 240
QY 240 HNFPAIMTKYISKPENIKLMMNLDRKSPNIQFEAFHVKFVVASPHKTQPIVEILLKNQ 299
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 HNFPAIMTKYISKPENIKLMMNLDRKSPNIQFEAFHVKFVVASPHKTQPIVEILLKNQ 300
QY 300 KLIEFLSSFKQKERTDDEQFADEXNYLIKQIRDLKTA 336
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 KLIEFLSSFKQKERTDDEQFADEXNYLIKQIRDLKTA 337

RESULT 5
ID MO25 DROME STANDARD; PRT; 339 AA.
AC P91891; Q9VV85;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MO25 protein (dMo25).
GN MO25 OR CG4083.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96268479; PubMed=8672247;
RA Nozaki M., Onishi Y., Togashi S., Miyamoto H.;
RT "Molecular characterization of the Drosophila Mo25 gene, which is
RL conserved among Drosophila, mouse, and yeast."
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Balow R.M., Bau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beremand B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

RESULT 6  
MO2M CAEEL

ID M02M\_CABEL STANDARD; PRT; 338 AA.  
 AC O18211;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical MO25-like protein Y53C12A.4 in chromosome II.  
 GN Y53C12A.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Beloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Kershaw J., Lennard N.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the Mo25 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z99277; CAB16486.1; -  
 DR PIR; T27129; T27129.  
 DR WormPep; Y53C12A.4; CE14890.  
 DR InterPro; IPR006938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 338 AA; 39431 MW; 1D0C343A5D9116F5 CRC64;  
 Query Match 59.1%; Score 1006.5; DB 1; Length 338;  
 Best Local Similarity 57.2%; Pred. No. 2.3e-57;  
 Matches 191; Conservative 60; Mismatches 78; Indels 5; Gaps 1;  
 QY 5 PLFSKSHKPAETVTKILKDNLAILE-KDKKTDKASEEYVSKSLQAMKEILCGTNEK 59  
 DB 4 PLFGGADKTPADVKNRLDALLVHGNTSRKVEKAEIEFAKMLAKTFIYGSAN 63  
 QY 60 EPTTEAVAQALYSSGLLVTLIADLIQIDFGKKDVTQIFNNILRQIGTRSPVVEYL 119  
 DB 64 EPNEQVTLAQEVYNAVLPMLIKHLKFEFECKDQVASFVNNLRRQIGTRSPVVEYL 123  
 QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKILFSNQFRDFFKYVELSTFDI 179  
 DB 124 AARPELITLLIGYEQPDIALTCGMLREAVRHEHLARIVLYSEYFORFVFVQSDVFDI 183  
 QY 180 ASDAPATFKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENVTYKRSKILGELILDR 239  
 DB 184 ATDAFTFKDLTKHKNCAEYLDNNYDRFFGYSALTNSENVYTRQSLKLLGELLDR 243  
 QY 240 HNFAMTKYISKPENILKLMNLRDKSPNIQEAHFVFKVVASPHKTOPIVEILLKNQOP 299  
 DB 244 HNFSTWANKYITSPENILKTMWELLDRKRNQIYEAHFVKIFVANPNKPIPTDILTRND 303  
 QY 300 KLIEFLSSFKERTDDEQFADEKNYLIKQIRDLK 333  
 DB 304 KLVEFUTAFHNDRTNDEQFNDEKAYLIKQIQELR 337  
 RESULT 7  
 YFV6\_SCHPO STANDARD; PRT; 329 AA.  
 ID YFV6\_SCHPO  
 AC Q9P708;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C1834.06c in chromosome I.  
 GN SPAC1834.06c.

OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,  
 RA Coiffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- SIMILARITY: Belongs to the Mo25 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL157734; CAB5774.1; -  
 DR PIR; T50117; T50117.  
 DR GenesDB Spombe; SPAC1834.06c; -  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 329 AA; 38521 MW; 073DD0607A64C952 CRC64;  
 Query Match 49.0%; Score 834.5; DB 1; Length 329;  
 Best Local Similarity 51.5%; Pred. No. 2e-46; 93; Indels 3; Gaps 2;  
 Matches 169; Conservative 63; Mismatches 63;  
 QY 6 PLFSKSHKPAETVTKILKDNLAILE-KDKKTDKASEEYVSKSLQAMKEILCGTNEKEPPT 64  
 DB 4 LFKRPKSTQDVVRCVCLDNLPLKLEINNDK--KSFEEVSKCLQNLRVSLCGTAIEVPDAD 61  
 QY 65 AVAQLAQELYSGLLVTLIADLIQIDFGKKDVTQIFNNILRQIGTRSPVVEYISAHPH 124  
 DB 62 LVSDLSFQIYQSNLPFLVRLVPLKLEFESKDTGLIFALLRRHVASRYPTVDYMLAHPQ 121  
 QY 125 ILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKILFSNQFRDFFKYVELSFEDIASAP 184  
 DB 122 IFFVLVSRYRQEVAFVAGTSGIURECSRHEALNVLNDRDFTWTFSLIQASSDFWASDAF 181  
 QY 185 ATFKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENVTYKRSKILGELILDRHFAI 244  
 DB 182 STFKSILLNKHKSQVAEFISYHFDEFFKQYTVLLKSENVYTKQSLKLLGELLNRSRV 241

```

QY 245 MTKYISKPNLKMMLLKDKNPNIQFEAFHVKFVSPHKTQPIVETILLKNQPKLIEF 304
Db 242 MTRYISSAENLKMMLLKDKNPNIQFEAFHVKFVSPHKTQPIVETILLKNQPKLIEF 301
QY 305 LGSFQKERTDDQFADEKKNVILKQIRDL 332
Db 302 LSAFHTRKNDQFNDFRAVINKQIERL 329

RESULT 8
ID DE76 CHLPR STANDARD; PRT; 321 AA.
AC Q9XFY6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Degreening related gene dee76 protein.
GN DEE76.
OS Chlorella protothecoides.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Auxenochlorella.
OX NCBI_TaxID=3075;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACC35;
RX MEDLINE=20256472; PubMed=10798614;
RA Hortensteiner S., Chinner J., Matile P., Thomas H., Donnison I.S.;
RT "Chlorophyll breakdown in Chlorella protothecoides: characterization
of degreening and cloning of degreening-related genes.";
RL Plant Mol. Biol. 42:439-450(2000).
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ238632; CAB42595.1; -.
DR InterPro: IPR008938; ARM.
DR InterPro: IPR004892; Mo25.
DR Pfam: PF03204; Mo25; 1.
SQ SEQUENCE 321 AA; 37262 MW; 918FD02964B09071 CRC64;

Query Match 45.5%; Score 776; DB 1; Length 321;
Best Local Similarity 52.0%; Pred. No. 1e-42;
Matches 156; Conservative 56; Mismatches 84; Indels 4; Gaps 3;

QY 32 DKTKDASEVSKLQAMKEILCGTNEKEPTTAVQAQELYSGLLVTLADLIQIDF 91
Db 19 ESKQDRVEDISKAISIKAEIAGEDEQSSSKSHAQGIASERCVRGLVSLVTLVTLDF 78
QY 92 EGKQDVQIENNLRRQI--GTSPTEVYISAPHPILFMLLKGVEAPOIALRCGIMLREC 149
Db 79 ETRKDVQVQIFCAIRITLEDGGR-PGRDYVLAHPDVLTSTPYGYDEPIALNCQGFREC 137
QY 150 IRHEPIAKIILFSGNQRFDFKQVELSTFDIASDAFAFFKDLLTRHKVLVADFTLEQNYDTI 209
Db 138 IRHEDIAKFVLECNLPFEELFEKLNQSVFEVSDAFATFKDLLTRHKQLVAAFLQENYEDF 197
QY 210 FEDYEKILQSENYVTKQSLKLGELILDRHNFAMTKYISKPNLKMMLLKDKNPNI 269
Db 198 FSQDKLLTSDNYVTRQSLKLGELILDRHNFAMTKYISKPNLKMMLLKDKNPNI 257
QY 270 QFEAFHVKFVSPHKTQPIVETILLKNQPKLIEFLSSFOKERTDDQFADEKKNVILKQI 329
Db 258 QFEAFHVKFVSPHKTQPIVETILLKNQPKLIEFLSSFOKERTDDQFADEKKNVILKQI 316

RESULT 9

```

```

MO2M ARATH
ID MO2M ARATH STANDARD; PRT; 343 AA.
AC Q9FGK3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical Mo25-like protein At5g47540.
GN At5g47540 OR MNJ7.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AB025628; BAB09080.1; -.
DR InterPro: IPR008938; ARM.
DR InterPro: IPR004892; Mo25.
DR Pfam: PF03204; Mo25; 1.
KW Hypothetical protein.
SQ SEQUENCE 343 AA; 39457 MW; 46950DCA9A82FEB5 CRC64;

Query Match 42.7%; Score 728; DB 1; Length 343;
Best Local Similarity 43.2%; Pred. No. 1.3e-39;
Matches 147; Conservative 79; Mismatches 100; Indels 14; Gaps 4;

QY 6 LFSKSHKNPAETIVKILKONLAILEK-----QDKTKDASEVSKLQAMKEILCGTNE 58
Db 4 LFKSKPTEADLVQRTRDLLLFSRSTSLPDLRDSKRDEKMAELSRNIRDMKSILYGNSE 63
QY 59 KEPPTTAVQAQELYSGLLVTLADLIQIDFEGKQDVQIENNLRRQIIGTRSPTEY 118
Db 64 ABEVAAEACQLTQEPFKEDTLRLTLCLPKNLTKDQATQVVALQROOVNSRLIASDY 123
QY 119 ISAPHPILFMLLKGVEAPOIALRCGIMLRECIRHEPIAKIILFSGNQRFDFKQVELSTFD 178
Db 124 LEANIDLMVLIEGFTMDALHYGAMFRECIRHQIVAKYVLESDHVKFFDYIQLPNFD 183
QY 179 IASDAFATFKDLLTRHKVLVADFTLEQNYDTIPEY--EKILQSENYVTKQSLKLGELIL 237
Db 184 IAAADAAATFKELLTRHKSTVAEFLTKNEDWFADYNSKLLSSNYITTRQAIKLGIDILL 243
QY 238 DRHNFAMTKYISKPNLKMMLLKDKNPNIQFEAFHVKFVSPHKTQPIVETILLKN 297
Db 244 DRSNAMVTKYSSRRNLRLMNLJRESKSIQIEAFHVKFVSPHKTQPIVETILLKN 303
QY 298 QPKLIEFLSSFOKERTDDQFADEKKNVILKQI-----RDL 332
Db 304 RSKLLRLADLKPDK-EDERFPAKSKQVLRLETAALPRDL 342

RESULT 10
MO2M ARATH
ID MO2M ARATH STANDARD; PRT; 343 AA.
AC Q9M0M4; O23570;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

```

RA RT RT RL RN RP RC RX RA RA RA RA RA RA RA RA RA RA RA RT RT RL CC CC CC CC

[illegible]

RA RT RT RL RN

RP

RX

RA RA

RA RA

RA RA

RA RA RA

RA RA

RA RA

RA  
PT

RT 2

CC BY

UU

UU

U U U

ॐ नमः

22

22

DR

DR DR

DR DR

DR KTW

SQ

5

E 2

or

22

၁၇

Qy

D**b** $\alpha\gamma$ 

Db

Qy

ሃገራችን

20

by

Dib

Qy

b6









QY 321 ---EKNYLKQIRDLKX 334  
Db 268 WDMQRPFTQEQLODFEE 284

RESULT 15  
AKA9 HUMAN  
ID AKA9 HUMAN STANDARD; PRT; 3911 AA.  
AC Q99996; O14869; O43355; O94895; Q9UQH3; Q9UQ04; Q9Y6B8; Q9Y6Y2;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE (PRKA9) A-kinase anchor protein 9 (Protein kinase A anchoring protein 9)  
DE protein 350 kDa (AKAP 350) (hgAKAP 350) (AKAP 120 like protein)  
DE (Hyperion protein) (Yotiao protein) (Centrosome- and Golgi-localized  
DE PKN-associated protein) (CG-NAP).  
GN AKAP9 OR AKAP450 OR AKAP350 OR KIAA0803.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 4).  
RC TISSUE=Brain;  
RX MEDLINE=98151389; PubMed=9482789;  
RA Lin J.W., Wysynski M., Madhavan R., Sealock R., Kim J.U., Sheng M.;  
RT "Yotiao, a novel protein of neuromuscular junction and brain that  
RT interacts with specific splice variants of NMDA receptor subunit  
RT NR1.";  
RL J. Neurosci. 18:2017-2027(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT GLN-1347 INS.  
RX MEDLINE=99219864; PubMed=10202149;  
RA Witzak O., Skaalheg B.S., Kerker G., Borners M., Tasken K.,  
RA Jahnsen T., Oerstavik S.;  
RT "Cloning and characterization of a cDNA encoding an A-kinase anchoring  
RT protein located in the centrosome, AKAP450.";  
RL EMBO J. 18:1858-1868(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC TISSUE=Brain;  
RX MEDLINE=99287934; PubMed=10358086;  
RA Takahashi M., Shibata H., Shimakawa M., Miyamoto M., Mukai H., Ono Y.;  
RT "Characterization of a novel grant scaffolding protein, CG-NAP, that  
RT anchors multiple signaling enzymes to centrosome and the Golgi  
RT apparatus.";  
RL J. Biol. Chem. 274:17267-17274(1999).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Kemmer W.A., Deiss S., Schwarz U.;  
RT "Cloning of Hyperion.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 323-3911 FROM N.A. (ISOFORM 2).  
RC TISSUE=Gastric parietal cell;  
RX MEDLINE=99115654; PubMed=9915845;  
RA Schmidt P.H., Dransfield D.T., Claudio J.O., Hawley R.G.,  
RA Trotter K.W., Milgram S.L., Goldenring J.R.;  
RT "AKAP350, a multiply spliced protein kinase A-anchoring protein  
RT associated with centrosomes.";  
RL J. Biol. Chem. 274:3055-3066(1999).  
RN [6]  
RP SEQUENCE OF 1802-3876 FROM N.A. (ISOFORM 5).  
RC TISSUE=Lymphoblast;  
RA Hinds K., Sutterer C., Becker M., Hawkins M.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 2157-3911 FROM N.A. (ISOFORM 6).  
RC TISSUE=lung;  
RA Milgram S.L., Goldenring J.R., Schmidt P.H.;  
RT "AKAP350: A multiply spliced family of proteins with centrosomal

RT association.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE OF 2212-3911 FROM N.A. (ISOFORM 2/3).  
RC TISSUE=Brain;  
RX MEDLINE=99087487; PubMed=9872452;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,  
RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XI.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:277-286(1998).  
RN [9]  
RP SEQUENCE OF 17-1800 FROM N.A.  
RL Wu X., Graves T., Bradshaw H.;  
RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Binds to type II regulatory subunits of protein kinase  
CC A. Scaffolding protein that assembles several protein kinases and  
CC phosphatases on centrosome and Golgi apparatus where physiological  
CC events can be regulated by phosphorylation state of protein  
CC substrates. Isoform 4/Yotiao is associated with the N-methyl-D-  
CC aspartate receptor and is specifically found in the neuromuscular  
CC junction (NMJ) as well as in neuronal synapses explaining that its  
CC role may be to organize postsynaptic specializations.  
CC -!- SUBUNIT: Interacts with the regulatory region of protein kinase N  
CC (PKN), protein phosphatase 2A (PP2A), protein phosphatase 1 (PPI)  
CC and the immature non-phosphorylated form of PKC epsilon.  
CC -!- SUBCELLULAR LOCATION: Centrosomal in many cell types and  
CC cytoplasmic in parietal cells.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=6;  
CC Name=1;  
CC IsoId=Q99996-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q99996-2; Sequence=VSP\_004102, VSP\_004107;  
CC Name=3; Synonyms=CG-NAP;  
CC IsoId=Q99996-3; Sequence=VSP\_004102, VSP\_004105, VSP\_004107;  
CC Name=4; Synonyms=Yotiao;  
CC IsoId=Q99996-4; Sequence=VSP\_004103, VSP\_004104;  
CC Name=5;  
CC IsoId=Q99996-5; Sequence=VSP\_004108;  
CC Name=6; Synonyms=AKAP350;  
CC IsoId=Q99996-6; Sequence=VSP\_004106, VSP\_004107, VSP\_004109;  
CC -!- TISSUE SPECIFICITY: Widely expressed. Isoform 4/Yotiao is highly  
CC expressed in skeletal muscle and in pancreas.  
CC -!- DOMAIN: RII-binding site, predicted to form an amphipathic helix,  
CC could participate in protein-protein interactions with a  
CC complementary surface on the R-subunit dimer.  
CC -!- CAUTION: Ref.6 sequence differs from that shown due to two  
CC frameshifts in positions 3782 and 3811.  
CC -!- CAUTION: Ref.9 sequence differs from that shown due to four  
CC frameshifts in positions 29, 1653, 1699 and 1735.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ131693; CAB407113.1; -  
DR EMBL; AB019691; BAB78718.1; -  
DR EMBL; AJ010770; CAA09361.1; -  
DR EMBL; AF026245; AAB86384.1; -  
DR EMBL; AF083037; AAD22767.1; -  
DR EMBL; AF004013; AAB96867.1; ALT\_FRAME.  
DR EMBL; AF017111; AAD39719.1; -  
DR EMBL; AB018346; BAA34523.1; -  
DR EMBL; AC000066; AAC60380.1; ALT\_FRAME.  
DR EMBL; HGNC:379; AKAP9.  
DR MIM; 604001; -  
DR GO; GO:0005813; C:centrosome; TAS.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:20:07 ; Search time 45 seconds  
(without alignments)  
2362.881 Million cell updates/sec

Title: US-10-025-730-1  
Perfect score: 1704  
Sequence: 1 MKKMPFSKGNPAEIVKI.....FADEKNYLIQIRDLKKTAP 337

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1462	85.8	289	4 Q96FG1	Q96fg1 homo sapien
2	1381	81.0	341	11 Q8VDZ8	Q8vdz8 mus musculu
3	1343	78.8	341	4 Q7ZAX0	Q7zax0 homo sapien
4	1288.5	75.6	343	13 Q803V8	Q803v8 brachydanio
5	1066.5	62.6	636	5 Q21643	Q21643 caenorhabdi
6	1063.5	62.4	377	5 Q722A5	Q722a5 caenorhabdi
7	875	51.3	205	11 Q8K312	Q8k312 mus musculu
8	671.5	39.4	345	10 Q8L9L9	Q8l9l9 arabidopsis
9	651.5	38.2	337	10 Q7XIQ4	Q7xiq4 oryza sativ
10	639.5	37.5	370	3 Q873K5	Q873k5 neurospora
11	590	34.6	322	10 Q8LIF3	Q8lif3 oryza sativ
12	435	25.5	103	11 Q8K038	Q8k038 mus musculu
13	300.5	17.6	509	5 Q7YVL6	Q7yy16 cryptospori
14	134.5	7.9	677	16 Q25188	Q25188 helicobacte
15	128	7.5	430	16 Q26049	Q26049 helicobacte
16	123.5	7.2	1285	16 Q9WXU3	Q9wxu3 thermotoga

17	119.5	7.0	1056	16 Q8REF7	Q8ref7 fusobacteri
18	119	7.0	1111	5 Q9VGE4	Q9vge4 drosophila
19	118.5	7.0	554	5 Q8IN90	Q8in90 drosophila
20	118.5	7.0	670	5 Q9VEC7	Q9vec7 drosophila
21	118.5	7.0	670	5 Q9NPM7	Q9nm7 drosophila
22	117	6.9	808	5 Q9GSH4	Q9gsh4 dictyosteli
23	117	6.9	808	5 Q8TI33	Q8ti33 dictyosteli
24	116.5	6.8	1135	5 Q9NJO4	Q9nj4 paramecium
25	116	6.8	911	16 Q8EUI7	Q8eui7 mycoplasma
26	116	6.8	1389	5 Q8I293	Q8i293 plasmodium
27	115.5	6.8	1111	5 Q9UOK5	Q9uok5 plasmodium
28	115.5	6.8	1946	5 Q97291	Q97291 plasmodium
29	115	6.7	473	11 Q8R436	Q8r436 mus musculu
30	115	6.7	2518	5 Q8IEH2	Q8ieh2 plasmodium
31	114.5	6.7	1941	5 Q8IAK6	Q8iak6 plasmodium
32	114	6.7	743	13 Q9YGE7	Q9yge7 oncorhynch
33	113.5	6.7	833	4 Q9UF54	Q9uf54 homo sapien
34	113.5	6.7	951	5 Q9VEC6	Q9vec6 drosophila
35	113.5	6.7	984	5 Q8IN89	Q8in89 drosophila
36	113.5	6.7	1335	4 Q86VH1	Q86vh1 homo sapien
37	113.5	6.7	1401	4 Q86VH2	Q86vh2 homo sapien
38	113	6.6	474	5 Q97233	Q97233 plasmodium
39	113	6.6	647	11 Q8CA10	Q8ca10 mus musculu
40	111.5	6.5	1925	5 Q8I2D1	Q8i2d1 plasmodium
41	111.5	6.5	2429	5 Q9VFB1	Q9vfb1 drosophila
42	111.5	6.5	2771	5 Q26216	Q26216 plasmodium
43	111	6.5	974	11 Q8WJ9	Q8wj9 mus musculu
44	111	6.5	2166	16 Q51465	Q51465 borrelia bu
45	111	6.5	2819	16 Q98QP8	Q98qp8 mycoplasma

## ALIGNMENTS

## RESULT 1

Q96FG1 ID Q96FG1 PRELIMINARY; PRT; 289 AA.  
AC DT Q96FG1; AC DT 01-DEC-2001 (Tremblrel. 19, Created)  
AT DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Strauberg R.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC010993; AAH10993.1; -  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 289 AA; 33738 MW; F57B9BFCF6ABF2D7 CRC64;

Query Match 85.8%; Score 1462; DB 4; Length 289;  
Best Local Similarity 99.7%; Pred. No. 1.6e-96;  
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 49 MKEILGCTNEKEPPTSAVAQLAQLYSSGLLVTLADLQIDPEGKDVDTQFNILRRQ 108  
Db 1 MKEILGCTNEKEPPTSAVAQLAQLYSSGLLVTLADLQIDPEGKDVDTQFNILRRQ 60  
Qy 109 IGRSTVEYISAHPHILFWLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFSNQFRDF 168  
Db 61 IGRSTVEYISAHPHILFWLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFSNQFRDF 120  
Qy 169 FKVELSTFDIASDAFATPKDLLTRHKVLVADPLEQNYDTIFEDYKLLQSENVYTKQS 228  
Db 121 FKVELSTFDIASDAFATPKDLLTRHKVLVADPLEQNYDTIFEDYKLLQSENVYTKQS 180

QY 229 LKLGELILDRHFAIMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQ 288  
 DB 181 LKLGELILDRHFAIMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQ 240  
 QY 289 PIVEILLKNQPKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 337  
 DB 241 PIVEILLKNQPKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 289  
 RESULT 2  
 Q8VDZ8 ID Q8VDZ8 PRELIMINARY; PRT; 341 AA.  
 AC Q8VDZ8; 01-MAR-2002 (T-EMBLrel. 20, Created)  
 DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE MO25 protein.  
 GN CAB39.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC020041; AAH20041.1; -  
 DR MGD; MGI:107436; Cab39.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 SQ SEQUENCE 341 AA; 39843 MW; E7FECAS59D6F811 CRC64;

Query Match 81.0%; Score 1381; DB 11; Length 341;  
 Best Local Similarity 81.0%; Pred. No. 1.2e-90;  
 Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;  
 QY 4 MPL-FSKSHKNPAEIVKILKONLAILEKO---DKKTDKASEEVSKSQAMKEILCGTNEK 59  
 DB 1 MPFPFGKSHKSPADIVKNIKESMAVLEKQDISDKAEKATEEVSKNVLVAMKEILYGTNEK 60  
 QY 60 EPPTTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 119  
 DB 61 EPQTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 120  
 QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFKXVELSTFDI 179  
 DB 121 CTQQNILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFKXVELSTFDI 180  
 QY 180 ASDAFATFKDLLTRHKLVAFLVADFLQNYDTIFEDYEKLQSENVTYKRSQSLKLLGELLDR 239  
 DB 181 ASDAFATFKDLLTRHKLVAFLVADFLQNYDTIFEDYEKLQSENVTYKRSQSLKLLGELLDR 240  
 QY 240 HNFATMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQPIVEILLKNQ 299  
 DB 241 HNFATMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQPIVEILLKNQ 300  
 QY 300 KLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 336  
 DB 301 KLIEFLSKFQNDRTDEQFNDEKTYLVKQIRDLKRAA 337

Query Match 75.6%; Score 1288.5; DB 13; Length 343;  
 Best Local Similarity 72.3%; Pred. No. 4.6e-84;  
 Matches 245; Conservative 49; Mismatches 40; Indels 5; Gaps 2;  
 QY 4 MPL-FSKSHKNPAEIVKILKONLAILEK---ODKTKDASEEVSKSQAMKEILCGTNE 58  
 DB 1 MPFPFGKSHKSPADIVKNIKESMAVLEKQDISDKAEKATEEVSKNVLVAMKEILYGTNE 60  
 QY 59 EPPTTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 118  
 DB 61 EPQTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 120  
 RESULT 3  
 Q724X0 ID Q724X0 PRELIMINARY; PRT; 341 AA.  
 AC Q724X0; 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE MO25-like protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhou Y., Yu L., Zhao S.Y.;  
 RT "Cloning of a new human cDNA homologous to MO25 gene."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF134480; AAP97257.1; -  
 SQ SEQUENCE 341 AA; 39774 MW; D3A60160E78C7A6D CRC64;  
 Query Match 78.8%; Score 1343; DB 4; Length 341;  
 Best Local Similarity 79.5%; Pred. No. 5.9e-88;  
 Matches 268; Conservative 31; Mismatches 34; Indels 4; Gaps 2;  
 QY 4 MPL-FSKSHKNPAEIVKILKONLAILEKO---DKKTDKASEEVSKSQAMKEILCGTNEK 59  
 DB 1 MPFPFGKSHKSPADIVKNIKESMAVLEKQDISDKAEKATEEVSKNVLVAMKEILYGTNEK 60  
 QY 60 EPPTTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 119  
 DB 61 EPQTEAVAQAQELYSGLLVTLIADQLIDFEGKKDVTOIFNNILRRQIGTRSPVVEYI 120  
 QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFKXVELSTFDI 179  
 DB 121 CTQQNILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFKXVELSTFDI 180  
 QY 180 ASDAFATFKDLLTRHKLVAFLVADFLQNYDTIFEDYEKLQSENVTYKRSQSLKLLGELLDR 239  
 DB 181 ASDAFATFKDLLTRHKLVAFLVADFLQNYDTIFEDYEKLQSENVTYKRSQSLKLLGELLDR 240  
 QY 240 HNFATMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQPIVEILLKNQ 299  
 DB 241 HNFATMTKYISKPENIKLMMNLLRDKSPNIOFEAFHVFKVVASPHKTQPIVEILLKNQ 300  
 QY 300 KLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 336  
 DB 301 KLIEFLSKFQNDRTDEQFNDEKTYLVKQIRDLKRAA 337  
 RESULT 4  
 Q803V8 ID Q803V8 PRELIMINARY; PRT; 343 AA.  
 AC Q803V8; 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Similar to MO25 protein.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC044172; AAH44172.1; -  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 SQ SEQUENCE 343 AA; 39820 MW; F10450DA0446268A CRC64;

QY 119 ISAHPHILFMLKGYEAPQIALRCGIMLRECIRHEPLAKILFNSQDRDFKXYVELSTFD 178  
 Db 121 ISSHQILFMLKGYETSEVALNCGMMLRECRHDPLARIVLFSDFYCFRYYEMSTFD 180  
 QY 179 IASDAFAFKDILLTRHKVILVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKILGELILD 238  
 Db 181 IASDAFAFKDILLTRHKIMCADFLETNYDVFTVEYKILHSENVYTKRQSLKILGELILD 240  
 QY 239 RHNFATKYSKIPENIKLMMNLLRDKSPNIOQFAFHVFKVFPVASHPKTQPIVEILKNQ 298  
 Db 241 RHNFATKYSKIPENIKLMMNLLRDKSPNIOQFAFHVFKVFPVASHPKTQPIVEILKNQ 298  
 QY 299 PKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 337  
 Db 301 SKLVEFLSHFQTRSEDEQFCDEKKNYLIKQIRDLKRPAP 339

## RESULT 5

Q21643 ID Q21643 PRELIMINARY; PRT; 636 AA.  
 AC Q21643  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN R02E12.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Leimbach D.;  
 RT "The sequence of C. elegans cosmid R02E12.";  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U53337; AAA96186.2; -.  
 DR PIR; T16650; T16650.  
 DR PIR; T16651; T16651.  
 DR WormPep; R02E12.2a; CR28410.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 636 AA; 72282 MW; 85D5853E9F0E3193 CRC64;

Query Match 62.6%; Score 1066.5; DB 5; Length 636;  
 Best Local Similarity 60.4%; Pred. No. 6.8e-68;  
 Matches 212; Conservative 53; Mismatches 69; Indels 17; Gaps 3;  
 QY 2 KQMP-LFSKSHKHPAIVIKLKNLAILEK-----QDKTKASEVSKSLQ 47  
 Db 258 KVMPLLFKSHKSPADVVKTLREVLTLDKLPPLKDKGNIQSKYDKALDEVSKNVA 317  
 QY 48 AMKEILCGTNEKEPPTPE---AVAQALQELYSGLLVTLADLQIDFEGKQVTFINNI 104  
 Db 318 MIKSFYIGNDSAEPSSEHVVQVQAQLAEVYNANILPMLIKMLPKFBECKDVQGFNNL 377

QY 105 LRQIGTRSPVTEYISAHPHILFMLKGYEAPQIALRCGIMLRECIRHEPLAKILFNSQ 164  
 Db 378 LRQIGTRSPVTEYISAHPHILFMLKGYEAPQIALRCGIMLRECIRHEPLAKILFNSQ 164  
 QY 165 FRDFFKVELSTFDIASDAFAFKDILLTRHKVILVADFLQNYDTIFEDYEKLQSENVT 224  
 Db 438 FYTFFLYVQSEVDFDSSDAFSTFKELTTRHKAIIAEFLDSNYDTFFAQYQNLNLSKNYV 497  
 QY 225 KRQSLKILGELILDHRNFATKYSKIPENIKLMMNLLRDKSPNIOQFAFHVFKVFPV 284  
 Db 498 KRQSLKILGELILDHRNFATKYSKIPENIKLMMNLLRDKSPNIOQFAFHVFKVFPV 284  
 QY 285 HKTQPIVEILLKNQPKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKT 335  
 Db 558 NKPKPISDILNRNREKLVEFLSEFHNDRDTDEQFDEKAYLIKQIOEMKSS 608

## RESULT 6

Q722A5 ID Q722A5 PRELIMINARY; PRT; 377 AA.  
 AC Q722A5  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein R02E12.2.  
 GN R02E12.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Wilson R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Leimbach D.;  
 RT "The sequence of C. elegans cosmid R02E12.";  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U53337; AAP40522.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 377 AA; 43470 MW; 8F9B8AA070216176 CRC64;

Query Match 62.4%; Score 1063.5; DB 5; Length 377;  
 Best Local Similarity 60.5%; Pred. No. 5.9e-68;  
 Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;  
 QY 4 MP-LFSKSHKHPAIVIKLKNLAILEK-----QDKTKASEVSKSLQAM 49  
 Db 1 MPELLFGKSHKSPADVVKTLREVLTLDKLPPLKDKGNIQSKYDKALDEVSKNVA 60  
 QY 50 KEILCGTNEKEPPTPE---AVAQALQELYSGLLVTLADLQIDFEGKQVTFINNI 106  
 Db 61 KSFYIGNDSAEPSSEHVVQVQAQLAEVYNANILPMLIKMLPKFBECKDVQGFNNL 120  
 QY 107 ROIGTRSPVTEYISAHPHILFMLKGYEAPQIALRCGIMLRECIRHEPLAKILFNSQ 166  
 Db 121 ROIGTRSPVTEYISAHPHILFMLKGYEAPQIALRCGIMLRECIRHEPLAKILFNSQ 166

```

QY 167 DFKYVELSTFIASDAFATFKDILLTRHKVLVADFLFQNVDTIPEDYEKLLQSENYYTKR 226
Db 181 TFFLYVQSEVDFISSDAFSTFKELTTRHKAIIAEFLDSNYDTFFPAQYNLLSNKYVYTR 240
QY 227 QSLKLGELIILDRHNFNFAIMTKYISKPNELKMMNLLRDKSPNIOFAHFVFKFVSPHK 286
Db 241 QSLKLGELIILDRHNFNFAIMTKYISKPNELKMMNLLRDKSPNIOFAHFVFKFVSPHK 300
QY 287 TQPIVEILLKNQPKLIFLSFQKERTDDQFADEKNYLIKQIRDLKKT 335
Db 301 PKPISDILNRNREKLVEFLSEFHNDRTDDEQFNDEKAYLIKQIQEMKSS 349

RESULT 7
Q8K312 PRELIMINARY; PRT; 205 AA.
AC Q8K312;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to calcium binding protein, 39 kDa (Fragment).
GN CAB39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
FT NON TER 1
SQ SEQUENCE 205 AA; 24582 MW; 015261A02F808169 CRC64;

Query Match 51.3%; Score 875; DB 11; Length 205;
Best Local Similarity 83.6%; Pred. No. 7.9e-55;
Matches 168; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 136 PQIARCGIMLRGICIRHEPLAKIILFSGNQFRDPFKYVELSTFIASDAFATFKDILLTRHK 195
Db 1 PEIALNCGIMLRGICIRHEPLAKIILWSEQVDFRYVEMSTFIASDAFATFKDILLTRHK 60
QY 196 VLVDLFQNVDTIPEDYEKLLQSENYYTKRQSLKLGELIILDRHNFNFAIMTKYISKPNEL 255
Db 61 LLSAEFLQHYDRFFSEYEKLLHSENVYTKRQSLKLGELIILDRHNFNFAIMTKYISKPNEL 120
QY 256 KLMNLLRDKSPNIOFAHFVFKFVSPHKTPQPIVEILLKNQPKLIFLSFQKERTDD 315
Db 121 KLMNLLRDKSPNIOFAHFVFKFVSPHKTPQPIVEILLKNQPKLIFLSFQKERTDD 180
QY 316 EQFADEKNYLIKQIRDLKKT 336
Db 181 EQFNDEKTYLVKQIRDLKRA 201

```

```

RESULT 8
Q8L9L9 PRELIMINARY; PRT; 345 AA.
AC Q8L9L9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RL annotation", 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY088359; AAM65898.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
KW Hypothetical protein.
SQ SEQUENCE 345 AA; 39841 MW; 2C46A3D3DEBB47AA CRC64;

Query Match 39.4%; Score 671.5; DB 10; Length 345;
Best Local Similarity 42.9%; Pred. No. 5e-40;
Matches 140; Conservative 68; Mismatches 113; Indels 5; Gaps 2;

QY 12 KNPAEIVKILKONLALILEKQD----KKTOKASEEVSLSQAMKEILCGTNEKEPPTFAVA 67
Db 12 KTFQEVVKAIRDSIMALDITKVVVEVKALEKEVEKNFSSLRGILSGOGETPNDQAV 71
QY 68 QLAQELYSGLVLTLIADLQILDFEGKQDVDTQTFNNILRQIRGTSPTVEYISAHPHILF 127
Db 72 QLALEFCKEDVSVSLVHKHLGWETRKOLLHCHWSILLKQKVGDTYCCVQYFEEHEFLD 131
QY 128 MLLKGEAPQIALRCGIMLRGICIRHEPLAKIILFSGNQFRDPFKYVELSTFIASDAFATF 187
Db 132 SLVVCYDNKEIALHCGSMLRECIKFPSPILAKYILLESACFELFFKFFVLPNFDVASFSTF 191
QY 188 KDLLTRHKVLVADFLFQNVDTIPEDYEKLLQSENYYTKRQSLKLGELIILDRHNFALMTK 247
Db 192 KDLLTRHDSVSVSEFLSHSTFEFDVYERLTSTSNVYTRKQSLKLSDFLEPPNSHIMKK 251
QY 248 YISKPNELKLMNLLRDKSPNIOFAHFVFKFVSPHKTPQPIVEILLKNQPKLIFLS 307
Db 252 FILEVRYLVKVIWTLAKSSKNIQISAFHIFKIFVANPNKQEVKVIILARNHEKLELLHD 311
QY 308 FQKER-TDDEQFADEKNYLIKQIRDL 332
Db 312 LSPGKGSEDDQPEEKELEIIEIQKL 337

RESULT 9
Q7XIQ4 PRELIMINARY; PRT; 337 AA.
AC Q7XIQ4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein OJ1316_A04.107.
GN OJ1316_A04.107.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RL clone:OJ1316_A04.107";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003822; BAC79608.1; -.
KW Hypothetical protein.
SQ SEQUENCE 337 AA; 38866 MW; B43138F9464720C8 CRC64;

```



```

Query Match      38.2%; Score 651.5; DB 10; Length 337;
Best Local Similarity 40.9%; Pred. No. 1.3e-38;
Matches 137; Conservative 76; Mismatches 113; Indels 9; Gaps 3;

QY 4 MFLFSKSHKMPA-----EIVKILKNLAILEKODKTD-KASEVSKSLQAMKEILCGTN 57
Db 1 MSFFFAASRPAPSPQELVRSIKESLLAL---DTRTGAKALEDEVKNVSTLRQTLSDG 57
QY 58 EKEPPTFAVAQLAQELYSGLLVTLADIQLIDFEGKVDVTFNNILRRQIGTRSPVTE 117
Db 58 EVPNQEQVLTALIEICKEDVLSLFVQNPWSLWEGRKDLAHCWSILLRQKVDYAYCCVQ 117
QY 118 YISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFRDFEYKVELSTF 177
Db 118 YIENHFDLLDFVVCYKNLEVALNCGNMLRECIKYPLAKIYLESSSEFFLQVLSNF 177
QY 178 DIASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLQSENIVTKRQSLKLLGLIL 237
Db 178 DIASDALNTEKDLITKHEAAVSEFLCSHYEQFFELYTRLLTSTNYVTRRQSVKFLSEFL 237
QY 238 DRHNFATMTKYSPENKLMNLLRDKSPNIQFEAFHVKVFAVSHPKTOPIVEILLKN 297
Db 238 EAPNAQIMKRYIIVESYLNIMIGLLKDTSKNIRICAFPHIFKVFVANENKPRDIIQVLVDN 297
QY 298 QPKLIEFLSSFKERTDDEQFADEKNYLIKOIRDL 332
Db 298 HRELLKILGNLPTSKGEDEQLEERDLIIEIKL 332

RESULT 10
Q873K5
ID Q873K5 PRELIMINARY; PRT; 370 AA.
AC Q873K5;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Probable protein required for conidiophore development.
GN B2314.090.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariales; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoehsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX284746; CAD70300.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 370 AA; 42812 MW; 2E82C63BD06D8B9E CRC64;

Query Match      37.5%; Score 639.5; DB 3; Length 370;
Best Local Similarity 37.2%; Pred. No. 1e-37;
Matches 134; Conservative 74; Mismatches 117; Indels 35; Gaps 4;

QY 8 SKSHKMPAETIKVLKNLAILEKODKTDKASEVSKSLQAMKEILCGTNEKEPPTFAVA 67
Db 8 ARRTNVSDLPRAREHV---KLDGPGQKVEZELAKVLSQMKLLQGHQEQEASLEHQY 64
QY 68 QLAQELYSGLLVTLADIQLIDFEGKVDVTFNNILRRQIGTRSPVTE-----YISA 121
Db 65 QLVTCMIEEDLLYLLAINLYRPLFDGRKDTQVIFSVFRFPNPARAEPLAVVVER 124
QY 122 HPILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQ----- 164
Db 125 RPQVLIELCKGYDHKESAQAGTAVLKLKSSPAATAVILHDDGDFGSSARGVGAIDRN 184

Query Match      34.6%; Score 590; DB 10; Length 322;
Best Local Similarity 38.5%; Pred. No. 3e-34;
Matches 129; Conservative 73; Mismatches 109; Indels 24; Gaps 4;

QY 4 MFLFSKSHKMPA-----EIVKILKNLAILEKODKTD-KASEVSKSLQAMKEILCGTN 57
Db 1 MSFFFAASRPAPSPQELVRSIKESLLAL---DTRTGAKALEDEVKNVSTLRQTLSDG 57
QY 58 EKEPPTFAVAQLAQELYSGLLVTLADIQLIDFEGKVDVTFNNILRRQIGTRSPVTE 117
Db 58 EVPNQEQVLTALIEICKEDVLSLFVQNPWSLWEGRKDLAHCWSILLRQKVDYAYCCVQ 117
QY 118 YISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFRDFEYKVELSTF 177
Db 118 YIENHFDLLDFVVCYKNLEVALNCGNMLRECIKYPLAKIYLESSSEFFLQVLSNF 177
QY 178 DIASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLQSENIVTKRQSLKLLGLIL 237
Db 178 DIASDALNTEKDLITKHEAAVSEFLCSHYEQFFELYTRLLTSTNYVTRRQSVKFLSEFL 237
QY 238 DRHNFATMTKYSPENKLMNLLRDKSPNIQFEAFHVKVFAVSHPKTOPIVEILLKN 297
Db 238 HRELLKILGNLPTSKGEDEQLEERDLIIEIKL 332

RESULT 11
Q8LIF3
ID Q8LIF3 PRELIMINARY; PRT; 322 AA.
AC Q8LIF3;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein (P0503D09.26 protein).
GN QJ1316.A04.9 OR P0503D09.26.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:QJ1316.A04.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0503D09.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003822; BAC06992.1; -.
DR EMBL; AP004555; BAC16736.1; -.
DR Gramene; Q8LIF3; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 37091 MW; 99434DFA7C2DCD21 CRC64;

Query Match      34.6%; Score 590; DB 10; Length 322;
Best Local Similarity 38.5%; Pred. No. 3e-34;
Matches 129; Conservative 73; Mismatches 109; Indels 24; Gaps 4;

QY 4 MFLFSKSHKMPA-----EIVKILKNLAILEKODKTD-KASEVSKSLQAMKEILCGTN 57
Db 1 MSFFFAASRPAPSPQELVRSIKESLLAL---DTRTGAKALEDEVKNVSTLRQTLSDG 57
QY 58 EKEPPTFAVAQLAQELYSGLLVTLADIQLIDFEGKVDVTFNNILRRQIGTRSPVTE 117
Db 58 EVPNQEQVLTALIEICKEDVLSLFVQNPWSLWEGRKDLAHCWSILLRQKVDYAYCCVQ 117
QY 118 YISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFRDFEYKVELSTF 177
Db 118 YIENHFDLLDFVVCYKNLEVALNCGNMLRECIKYPLAKIYLESSSEFFLQVLSNF 177
QY 178 DIASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLQSENIVTKRQSLKLLGLIL 237
Db 178 DIASDALNTEKDLITKHEAAVSEFLCSHYEQFFELYTRLLTSTNYVTRRQSVKFLSEFL 237
QY 238 DRHNFATMTKYSPENKLMNLLRDKSPNIQFEAFHVKVFAVSHPKTOPIVEILLKN 297
Db 238 HRELLKILGNLPTSKGEDEQLEERDLIIEIKL 332

```



DR SMART; SM00436; TOP1BC; 1.  
DR SMART; SM00493; TOP1RM; 1.  
KW Hypothetical protein; Isomerase; Complete proteome.  
SQ SEQUENCE 677 AA; 77677 MW; 43285B81F1092BB4 CRC64;

Query Match 7.9%; Score 134.5; DB 16; Length 677;  
Best local similarity 21.6%; Pred. No. 0.24;  
Matches 88; Conservative 59; Mismatches 134; Indels 127; Gaps 16;

QY 7 FSKSHKNPA-EIVKILKDNL-----AILEKQDKK---TDKASEEVSLSQAMKE 51  
DB 222 FKFKDNEASQFLKDLKGLGMSVLVSLKESLSNKEPKPFTTSKLLSQASKSLKI--- 278  
QY 52 ILCGTNEKEPPTRAVAQAQLYSSGLLVTLIADQLIDFEGKQDVTOIFNNILRRQIGT 111  
DB 279 -----PTKEIAQAQLKFEAGLITVHTDSEFLSPYLKHEHVFPEIY----- 322  
QY 112 RSTTV-----EYIS-----AHPHILFMLLKGVEAPQIALRCGIMLRCEIRHE 153  
DB 323 --PSVQYREYKAGKNSQAEAEAIRITHPHALKDLKGVCDAKISEELALKLYLIYTN 380  
QY 154 PL---AKILLFSQFRDFKYYVELSTEDIASDAFATFKDLLTRHKVLVADFLSQNDTIF 210  
DB 381 TICSQRRNALY-NOYDCIFK-----IKSEFKLJSFKLLKKGFLIEBELIQGKEEIN 431  
QY 211 EDYEKLLQSENYVTKQSLKLLGELILDRHNFATMTKYISKPENLKLMMNLLLRDKSPNIQ 270  
DB 432 RE-EQSEIEENFSLKENDSVLPKEVFIKK-----IEKPSPKPYKESAFIFLLSESG--- 481  
QY 271 FEAFHFKVFPVSPHKTQPIVEILLKNQ-----PKLIEFLSSFOKERTDD- 315  
DB 482 -----IGRPSTYASFLDLLLKKYISIDTKTNAITPSTQGLEVISFFKKDKVEVDF 531  
QY 316 -----EQF-----ADEKNYLKIQIRDLKKTA 336  
DB 532 IALTSDKSKLGNITTKQFECLELIMRGEASYKFMLEVISLKLKSTA 578

RESULT 15  
Q26049  
ID Q26049 PRELIMINARY; PRT; 430 AA.  
AC Q26049;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein Hp1520.  
GN Hp1520.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori.";  
RL Nature 388:539-547 (1997).  
DR ENBL; AE000650; AAD08565.1; --  
DR PIR; H64709; H64709.  
DR TIGR; HPI520; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 430 AA; 50573 MW; 23DC6FE5E956B629 CRC64;

Query Match 7.5%; Score 128; DB 16; Length 430;  
Best local similarity 20.9%; Pred. No. 0.39;  
Matches 82; Conservative 73; Mismatches 135; Indels 102; Gaps 20;

QY 7 FSKSHKNPAEII---VKILKDNLAILEKQDKKTKTDKASEEVSLSQAMKEILCGTNEKEPPT 62  
DB 60 FYENRKSKEIEFNGEKILKENVAVFHSYDE--EFSSEDSVTTTFMAKSDI-----KQY 111  
QY 63 TEAVAQAQLYSSGLLVTL--IA-----DLQLIDFEGKQDVTOIFNNILR----- 106  
DB 112 DNTILLEKEE--KKALLKSLRDIASGFDYEEIEIKTKNEKNKSFYIELDNHLTEIESSEK 169  
QY 107 -----ROIGTRSPTV-EYISAHPHILFMLLKGVEAPQIALRCGIMLRCEIRHEPLAKII 159  
DB 170 HYSFKYRDFDGSKKVDFVNGKHDLIEQYFNKYQ-----ELLSQSK 211  
QY 160 LF-----SNQPRDFKYYVELSTEDIASDAFATFKDLLTRHKVLVADFLSQ----- 204  
DB 212 IFKHMNSGDFGTNHADDLKKALENNRFFKANHSLSKIAGEBITNYQKL-SDIFENKNRIL 270  
QY 205 NYDTIEDEYKLLQSENYVTKQSLKLLGELI-----LDRHNF--AINTKYISKP 252  
DB 271 NNSBELSKFDKI---EKVINANKELKAFKADAIKONTLLTEFLDYDSFRKKLVLSYLKQV 327  
QY 253 -ENLKLMMNLLLRDKSPNIQFEAFHFKVFPVSPHKTQPIVEILLKNQPKLIEFLSSFOKE 311  
DB 328 IQNVKSLVNLVREKKPEIE-----EIIKQASKDQKQEWESVIEIF--NORFLVPPFKVELQ 361  
QY 312 R-----TDDEQ-----FADEKNYLKIQIRDLKK 334  
DB 382 KDILLNKDAARFRFIFSDDDQDMNVQKEDLQK 413

Search completed: April 12, 2004, 10:28:39  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 09:47:16 ; Search time 59 Seconds  
(without alignments)  
1613.873 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 1704

Sequence: 1 MKKMPLEFSKHNPAEIVK.....FADEKNYLIKQIRDLKKTAP 337

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	100.0	337	3	AAY94247 Human cal
2	1704	100.0	337	4	AAB82090 Human acu
3	1704	100.0	337	4	AAM39078 Human pol
4	1466	86.0	289	4	AAB94139 Human pro
5	1381	81.0	341	4	AAB48970 Human ANI
6	1381	81.0	496	4	Aae10858 Gal4-huma
7	1381	81.0	552	4	Aae10859 LexA-huma
8	1376	80.8	341	3	AAY94248 Mouse cal
9	1354	79.5	354	4	ABG23844 Novel hum
10	1297.5	76.1	350	4	AAB20387 Human acu
11	1162	68.2	237	4	AAM40864 Human pol
12	1111	65.2	339	4	ABB60392 Drosophil
13	1109	65.1	339	3	AAY94249
14	1063.5	62.4	377	3	AAY94250 C. elegans
15	716.5	42.0	343	3	AAG45273 Arabidops
16	689.5	40.5	300	3	AAG23886 Arabidops
17	685.5	40.2	400	3	AAG51052 Arabidops
18	685.5	40.2	504	3	AAG51051 Arabidops
19	685	40.2	300	3	AAG45274 Arabidops
20	685	40.2	300	3	AAG30714 Arabidops
21	685	40.2	305	3	AAG30713 Arabidops
22	684.5	40.2	326	3	AAG51053 Arabidops
23	675.5	39.6	290	3	AAG23887 Arabidops
24	671.5	39.4	345	3	AAG50508 Arabidops
25	638.5	37.5	320	3	AAG50509 Arabidops

## ALIGNMENTS

### RESULT 1

AAY94247

ID AAY94247 standard; protein; 337 AA.

XX AC AAY94247;

XX AC AC (first entry)

DT 10-AUG-2000

XX Human calcium binding protein hCBP.

DE Human; calcium binding protein; cancer; inflammation; CBP;

KW reproductive disorder; autoimmune disorder; developmental disorder;

KW seizure disorder; immune disorder; infection.

XX Homo sapiens.

XX WO200029580-A1.

XX 25-MAY-2000.

PD 12-NOV-1999; 99WO-US027027.

XX 13-NOV-1998; 98US-00190965.

XX (INCY-) INCYTE PHARM INC.

PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX WPI; 2000-387793/33.

DR N-PSDB; AAA27332.

XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.,

PT diagnosis, prevention and treatment of cancers, immune, developmental or

PT reproductive disorders.

XX Claim 1; Fig 1; 72pp; English.

XX The present sequence is the human calcium binding protein hCBP. It was

CC obtained by screening a coronary artery smooth muscle cDNA library, from

CC which five overlapping nucleic acids were isolated, sequenced and

CC expressed to give the protein. The protein and the gene encoding it are

CC useful for the diagnosis and treatment of the following types of

CC disorder: cancers (such as adenocarcinomas), reproductive disorders (such

CC as infertility, ovulatory defects, endometriosis, disruptions of the

CC oestrus and menstrual cycles, polycystic ovary syndrome and ovarian

CC hyperstimulation), autoimmune disorders (such as benign prostatic

CC hyperplasia and prostatitis), developmental disorders (such as Cushing's

CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary

Aag23888 Arabidops  
Aag30715 Arabidops  
Aag45275 Arabidops  
Aag50591 Arabidops  
Aag41151 Zea mays  
Aag41152 Zea mays  
Aag41153 Zea mays  
Aag20291 Human ORF  
Abg25372 Novel hum  
Abp34081 Human ORF  
Abg23843 Novel hum  
Abr92048 Human cer  
Abr92047 Human cer  
Adb57140 Human A-k  
Abr92050 Human cer  
Abb30817 Peptide #  
Abg38772 Human pep  
Abu25521 Protein e  
Abg70293 Human nov

26 539.5 31.7 213 3 AAG23888  
27 533 31.3 213 3 AAG30715  
28 533 31.3 213 3 AAG45275  
29 478.5 28.1 197 3 AAG50591  
30 467.5 27.4 154 3 AAG41151  
31 453.5 26.6 148 3 AAG41152  
32 438.5 25.7 139 3 AAG41153  
33 250.5 14.7 236 5 ABP02921  
34 241 14.1 639 4 ABG25372  
35 227.5 13.4 135 5 ABP34081  
36 226.5 13.3 383 4 ABG23843  
37 128.5 7.5 3899 6 ABR92048  
38 128.5 7.5 3907 6 ABR92047  
39 128.5 7.5 3911 7 ADB57140  
40 128.5 7.5 3917 6 ABR92050  
41 128.5 7.5 3925 6 ABR92049  
42 125 7.3 660 4 ABB30817  
43 125 7.3 660 5 ABG38772  
44 120.5 7.1 1182 6 ABU25521  
45 117.5 6.9 709 5 ABG70293

CC neuropathies, seizure disorders, immune disorders (such as AIDS,  
 CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's  
 CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,  
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative  
 CC colitis), and viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections  
 XX  
 XX SQ Sequence 337 AA;  
 Query Match 100.0%; Score 1704; DB 3; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-146;  
 Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKMPFLSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQAMKEILCGTNEKE 60  
 DB 1 MKKMPFLSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQAMKEILCGTNEKE 60  
 QY 61 PPTTEVAQAQLAQLYSSGLLVTLIADQLIDFEGKKDVTQIFNNILRRQIGTRSPVVEYIS 120  
 DB 61 PPTTEVAQAQLAQLYSSGLLVTLIADQLIDFEGKKDVTQIFNNILRRQIGTRSPVVEYIS 120  
 QY 121 AHPHILFMLLKGYEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYYVELSTFDIA 180  
 DB 121 AHPHILFMLLKGYEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYYVELSTFDIA 180  
 QY 181 SDAPATFKDLLTRHKVLVADLEQNYDTIFEDYEKLLQSENVTYKRSKLIGELILDRH 240  
 DB 181 SDAPATFKDLLTRHKVLVADLEQNYDTIFEDYEKLLQSENVTYKRSKLIGELILDRH 240  
 QY 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTOPIVEILLKNQPK 300  
 DB 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTOPIVEILLKNQPK 300  
 QY 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337  
 DB 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337

## RESULT 2

AAB82090  
 ID AAB82090 standard; protein; 337 AA.  
 AC AAB82090;  
 DT 26-JUN-2001 (first entry)  
 XX Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP.  
 XX Human; cerebroprotective; neuroprotective; vulnerary; vaccine;  
 KW gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;  
 KW stroke; acute head trauma; multiple sclerosis; spinal cord injury.  
 XX  
 OS Homo sapiens.  
 XX WO200123552-A1.  
 XX N-PSDB; AAF86462.  
 PD 05-APR-2001.  
 XX  
 PF 18-SEP-2000; 2000WO-EP009132.  
 XX  
 PR 24-SEP-1999; 99EP-00118848.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 XX  
 PI Den Daas I, Duecker K;  
 XX  
 XX WPI; 2001-308142/32.  
 DR N-PSDB; AAF86462.  
 XX  
 XX Novel human acute neuronal induced calcium binding polypeptide, and  
 PT polynucleotides encoding them useful for diagnosing or treating stroke,  
 PT acute head trauma, multiple sclerosis and spinal cord injury.  
 XX

PS Claim 1; Page 41-42; 45pp; English.  
 XX  
 CC The present sequence is the protein sequence for human Acute Neuronal  
 CC Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and  
 CC protein are useful for treating stroke, acute head trauma, multiple  
 CC sclerosis and spinal cord injury. ANIC-BP coding sequence and protein are  
 CC also useful as vaccines for inducing an immunological response in a  
 CC mammal  
 XX  
 XX SQ Sequence 337 AA;

Query Match 100.0%; Score 1704; DB 4; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-146;  
 Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKMPFLSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQAMKEILCGTNEKE 60  
 DB 1 MKKMPFLSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQAMKEILCGTNEKE 60  
 QY 61 PPTTEVAQAQLAQLYSSGLLVTLIADQLIDFEGKKDVTQIFNNILRRQIGTRSPVVEYIS 120  
 DB 61 PPTTEVAQAQLAQLYSSGLLVTLIADQLIDFEGKKDVTQIFNNILRRQIGTRSPVVEYIS 120  
 QY 121 AHPHILFMLLKGYEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYYVELSTFDIA 180  
 DB 121 AHPHILFMLLKGYEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYYVELSTFDIA 180  
 QY 181 SDAPATFKDLLTRHKVLVADLEQNYDTIFEDYEKLLQSENVTYKRSKLIGELILDRH 240  
 DB 181 SDAPATFKDLLTRHKVLVADLEQNYDTIFEDYEKLLQSENVTYKRSKLIGELILDRH 240  
 QY 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTOPIVEILLKNQPK 300  
 DB 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTOPIVEILLKNQPK 300  
 QY 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337  
 DB 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337

## RESULT 3

AAM39078  
 ID AAM39078 standard; protein; 337 AA.  
 AC AAM39078;  
 DT 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 2223.  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX WO200153312-A1.  
 PN  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-0052317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.  
 PA (HYPE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J, Zhao Q;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AA158234.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX Example 4; SEQ ID NO 2223; 10078pp; English.  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
 CC encoded polypeptides (AA138642-AA142213) with nootropic,  
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX Sequence 337 AA;  
 SQ  
 Query Match 100.0%; Score 1704; DB 4; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-146;  
 Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKQWPLFSKSHKNPABIVKILKDNLAILEKQDKKTKASBEVSKSLQAMKEILCGTNEKE 60  
 DB 1 MKQWPLFSKSHKNPABIVKILKDNLAILEKQDKKTKASBEVSKSLQAMKEILCGTNEKE 60  
 QY 61 PPTVAQAQAEYSSGLLVTLADIQLIDFEGKDVQTFNNILRRQIGTRSPVVEYS 120  
 DB 61 PPTVAQAQAEYSSGLLVTLADIQLIDFEGKDVQTFNNILRRQIGTRSPVVEYS 120  
 QY 121 AHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFPKYVELSTFDIA 180  
 DB 121 AHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDFPKYVELSTFDIA 180  
 QY 181 SDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLLQSNYVTKRQSLKLGELLDRH 240  
 DB 181 SDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLLQSNYVTKRQSLKLGELLDRH 240  
 QY 241 NFAIMTKYISKPNLKMNNLLRDKSPNIQFEAFHVKFVSPHKTQPIVEILLKNQPK 300  
 DB 241 NFAIMTKYISKPNLKMNNLLRDKSPNIQFEAFHVKFVSPHKTQPIVEILLKNQPK 300  
 QY 301 LIEFLSSFOKERTDDEQFADEKKNYLKQIRDLLKKTAP 337  
 DB 301 LIEFLSSFOKERTDDEQFADEKKNYLKQIRDLLKKTAP 337  
 RESULT 4  
 ID AAB94139  
 XX AAB94139 standard; protein; 289 AA.  
 AC AAB94139;  
 XX 26-JUN-2001 (first entry)  
 DT Human protein sequence SEQ ID NO:14408.  
 DE  
 XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 OS  
 XX EP1074617-A2.  
 PN  
 XX 07-FEB-2001.  
 PD  
 XX 28-JUL-2000; 2000EP-00116126.  
 PF  
 XX 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 DR  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX Claim 8; SEQ ID NO 14408; 2537pp + Sequence Listing; English.  
 PS  
 XX The present invention describes primer sets for synthesising 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 XX Sequence 289 AA;  
 SQ  
 Query Match 86.0%; Score 1466; DB 4; Length 289;  
 Best Local Similarity 99.7%; Pred. No. 1e-124;  
 Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 49 MKEILCGTNEKEPPTVAQAQAEYSSGLLVTLADIQLIDFEGKDVQTFNNILRRQ 108  
 DB 1 MKEILCGTNEKEPPTVAQAQAEYSSGLLVTLADIQLIDFEGKDVQTFNNILRRQ 60  
 QY 109 ICTRSPPTVEYISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDF 168  
 DB 61 ICTRSPPTVEYISAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQPRDF 120  
 QY 169 FKVELSTFDIASDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLLQSNYVTKRQS 228  
 DB 121 FKVELSTFDIASDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLLQSNYVTKRQS 180  
 QY 229 LKILGELLDRHNFAMTKYISKPNLKMNNLLRDKSPNIQFEAFHVKFVSPHKTQ 288  
 XX





Query Match 81.0%; Score 1381; DB 4; Length 496;  
 Best Local Similarity 81.0%; Pred. No. 1.1e-116;  
 Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;

QY 4 MPL-FSKSHKNPAEIVKILKDNLAILEKQ---DKKTDKASEEVSYSLOAMKEILCGTNEK 59  
 DB 156 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 215  
 QY 60 EPPEEVAQAQELYSGLVTLADLQIDFEGKKDVTOIFNNILRQIGTRSPVVEYI 119  
 DB 216 EPQTEAVAQAQELYNGLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVVEYI 275  
 QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQPRDFKYYVELSTFDI 179  
 DB 276 CTQONILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQPRDFKYYVELSTFDI 335  
 QY 180 ASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKRQSLKLLGELLDR 239  
 DB 336 ASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKRQSLKLLGELLDR 395  
 QY 240 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVVFVSPHKTQPIVILLKNOP 299  
 DB 396 HNFATMTKYISKPENLKMNNLLRDKSRNIOFEAFHVKVVFVSPHKTQPIVILLKNQA 455  
 QY 300 KLIEFLSSFKQRTDDEQFADEKNYLIKQIRDLKKA 336  
 DB 456 KLIEFLSKFQNDRTDEQFNDKTYLVKQIRDLKKA 492

RESULT 7  
 AAEE10859  
 ID AAEE10859 standard; protein; 552 AA.  
 AC AAEE10859;  
 XX  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX

LexA-human ANIC-BP-1 fusion protein.  
 Human; acute neuronal induced calcium binding protein type 1 ligand;  
 ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;  
 Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;  
 gene therapy; fusion protein; LexA protein.  
 OS Homo sapiens.  
 OS Unidentified.  
 OS Chimeric.

Key Location/Qualifiers  
 FH Region 1..202  
 FT /note="LexA protein"  
 FT 203..552  
 FT /note="Human ANIC-BP-1 protein"  
 XX  
 XX WO200170771-A2.  
 XX  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX  
 PF 20-MAR-2001; 2001WO-EF003149.  
 XX  
 XX  
 PR 21-MAR-2000; 2000EP-00106110.  
 XX  
 XX (MERE ) MERCK PATENT GMBH.  
 XX  
 XX Den Daas I, Duecker K, Hock B;  
 XX  
 XX WPI; 2001-607519/69.  
 XX  
 XX Novel acute neuronal induced calcium binding protein type 1 ligand  
 PT polypeptides, useful in the treatment of stroke, head trauma, multiple  
 PT sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord  
 PT injury.  
 XX

PS Disclosure; Page 44-46; 46pp; English.  
 XX  
 CC The invention relates to human acute neuronal induced calcium binding  
 CC protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.  
 CC Sequences of the invention are useful for treating human diseases  
 CC including stroke, head trauma, multiple sclerosis, Parkinson's disease,  
 CC Alzheimer's disease and spinal cord injury. They are also useful as  
 CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound  
 CC soluble receptors. Polynucleotides of the invention are useful as  
 CC diagnostic reagents, for chromosome localization studies, and as valuable  
 CC tools for tissue expression studies. They are also useful in gene  
 CC therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein  
 CC comprising the LexA protein and a C-terminally linked human ANIC-BP-1  
 CC protein  
 XX  
 SQ Sequence 552 AA;

Query Match 81.0%; Score 1381; DB 4; Length 552;  
 Best Local Similarity 81.0%; Pred. No. 1.3e-116;  
 Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;

QY 4 MPL-FSKSHKNPAEIVKILKDNLAILEKQ---DKKTDKASEEVSYSLOAMKEILCGTNEK 59  
 DB 212 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 271  
 QY 60 EPPEEVAQAQELYSGLVTLADLQIDFEGKKDVTOIFNNILRQIGTRSPVVEYI 119  
 DB 272 EPQTEAVAQAQELYNGLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVVEYI 331  
 QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQPRDFKYYVELSTFDI 179  
 DB 332 CTQONILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFNSQPRDFKYYVELSTFDI 391  
 QY 180 ASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKRQSLKLLGELLDR 239  
 DB 392 ASDAFATFKDILLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKRQSLKLLGELLDR 451  
 QY 240 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVVFVSPHKTQPIVILLKNOP 299  
 DB 452 HNFATMTKYISKPENLKMNNLLRDKSRNIOFEAFHVKVVFVSPHKTQPIVILLKNQA 511  
 QY 300 KLIEFLSSFKQRTDDEQFADEKNYLIKQIRDLKKA 336  
 DB 512 KLIEFLSKFQNDRTDEQFNDKTYLVKQIRDLKKA 548

RESULT 8  
 AAAY94248  
 ID AAAY94248 standard; protein; 341 AA.  
 XX  
 AC AAAY94248;  
 XX  
 XX  
 DT 10-AUG-2000 (first entry)  
 XX  
 XX Mouse calcium binding protein MO25.  
 DE  
 DE  
 DE  
 KW Mouse; calcium binding protein; cancer; inflammation; MO25; CBP;  
 KW reproductive disorder; autoimmune disorder; developmental disorder;  
 KW seizure disorder; immune disorder; infection.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200029580-A1.  
 XX  
 XX  
 PD 25-MAY-2000.  
 XX  
 XX  
 PF 12-NOV-1999; 99WO-US027027.  
 XX  
 XX  
 PR 13-NOV-1998; 98US-00190965.  
 XX  
 XX (INCY-) INCYTE PHARM, INC.  
 XX  
 XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX WPI; 2000-387793/33.  
 XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
 PT diagnosis, prevention and treatment of cancers, immune, developmental or  
 PT reproductive disorders.  
 XX  
 XX Disclosure; Page 66-67; 72pp; English.  
 XX  
 CC The present sequence is the mouse calcium binding protein MO25. It was  
 CC used in a sequence alignment to identify human calcium binding protein  
 CC hCBP. The hCBP protein and the gene encoding it are useful for the  
 CC diagnosis and treatment of the following types of disorder: cancers (such  
 CC as adenocarcinomas), reproductive disorders (such as infertility,  
 CC ovulatory defects, endometriosis, disruptions of the estrus and  
 CC menstrual cycles, polycystic ovary syndrome and ovarian  
 CC hyperstimulation), autoimmune disorders (such as benign prostatic  
 CC hyperplasia and prostatitis), developmental disorders (such as Cushing's  
 CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary  
 CC neuropathies, seizure disorders, immune disorders (such as AIDS,  
 CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's  
 CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,  
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative  
 CC colitis), and viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections  
 XX  
 SQ Sequence 341 AA;

Query Match 80.8%; Score 1376; DB 3; Length 341;  
 Best Local Similarity 80.7%; Pred. No. 2e-116;  
 Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;  
 Qy 4 MPL-FSKSHKPAETIVKILKONLAILEKQ---DKKTDKASEEVSLSLOAMKEILCGTNEK 59  
 Db 1 MPFFGSKHSPADIVKLNKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60  
 Qy 60 EPTEAVAQAQLYSSGLVTLIADLQIDFEGKKDVTQIFNNILRRQIGTRSPVVEYI 119  
 Db 61 EPQTEAVAQAQLYSSGLVTLIADLQIDFEGKKDVAQIFNNILRRQIGTRPTVEYI 120  
 Qy 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFDFPKYVELSTFDI 179  
 Db 121 CTQONILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFDFPKYVELSTFDI 180  
 Qy 180 ASDAFATPKDLTRHKVLVADFLQNYDTIFEDYEKLQSENVYTKQSLKLGELLDR 239  
 Db 181 ASDAFATPKDLTRHKVLVADFLQNYDTIFEDYEKLQSENVYTKQSLKLGELLDR 240  
 Qy 240 HNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAFHFKVFAVSPHKTQPIVEILLKNQ 299  
 Db 241 HNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAFHFKVFAVSPHKTQPIVEILLKNQ 300  
 Qy 300 KLIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKTKA 336  
 Db 301 KLIEFLSKFQNDRTDEQFNDKTYLVKQIRDLKRAA 337

RESULT 9  
 ABG23844  
 ID ABG23844 standard; protein; 354 AA.  
 XX  
 AC ABG23844;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #23835.  
 XX  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 FN

XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 PF  
 XX 31-MAR-2000; 2000US-00540217.  
 PR  
 XX 23-AUG-2000; 2000US-00649167.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 FA  
 XX Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI: 2001-639362/73.  
 PI  
 XX N-PSDB; AAS88031.  
 DR  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 54203; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridization probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 354 AA;  
 Query Match 79.5%; Score 1354; DB 4; Length 354;  
 Best Local Similarity 79.2%; Pred. No. 2.1e-114;  
 Matches 267; Conservative 33; Mismatches 33; Indels 4; Gaps 2;  
 Qy 4 MPL-FSKSHKPAETIVKILKONLAILEKQ---DKKTDKASEEVSLSLOAMKEILCGTNEK 59  
 Db 14 MPFFGSKHSPADIVKLNKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 73  
 Qy 60 EPTEAVAQAQLYSSGLVTLIADLQIDFEGKKDVTQIFNNILRRQIGTRSPVVEYI 119  
 Db 74 DPQTEAGAQAQLYSSGLVTLIADLQIDFEGKKDVAQIFNNILRRQIGTRPTVEYI 133  
 Qy 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFDFPKYVELSTFDI 179  
 Db 134 CTQONILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQFDFPKYVELSTFDI 193  
 Qy 180 ASDAFATPKDLTRHKVLVADFLQNYDTIFEDYEKLQSENVYTKQSLKLGELLDR 239  
 Db 194 ASDAFATPKDLTRHKVLVADFLQNYDTIFEDYEKLQSENVYTKQSLKLGELLDR 253  
 Qy 240 HNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAFHFKVFAVSPHKTQPIVEILLKNQ 299  
 Db 254 HNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAFHFKVFAVSPHKTQPIVEILLKNQ 313  
 Qy 300 KLIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKTKA 336  
 Db 314 KLIEFLSKFQNDRTDEQFNDKTYLVKQIRDLKRAA 350

QY	120	SAHPHILFMLLKGYEAPQIALRCGIMLR	RECIRHEPLAKILILSNQRPDRFKVVELSTEDI	179
Db	121	CTQCNILFMLLKGYEPEIALNCGIMLR	RECIRHEPLAKILILSEQDYDFRIVEMSTEDI	180
QY	180	ASDAFAFVKOLLTHKXVYLVADFL	EQNYDITFDEYKLLQSENVTXKQSLKLLGELLDR	239
Db	181	ASDAFAFVKOLLTHKLLSAEFL	EQHYDRFFSEYKLLHSENVYVTKQSLKLLGELLDR	240
QY	240	HNFAIMTKYISKPENLKLMMNLLR	DKSPNIQFAFHVFKVFAVSPHKTOPIVEILLKNQP	299
Db	241	HNFTIMTKYISKPENLKLMMNLLR	DKSRNIQFAFHVFKVFAVSPHKTOPIVDILLKNQA	300
QY	300	KLIFLGLSFQKERTD-----	DEQFADEKNYLKIQIRDLKKTA	336
Db	301	KLIFLGLSKFQNDRTDCMS	SVPTNSRVDLRVKPRTRGIRDLKRPA	346
RESULT 11				
AAM40864				
ID	AAM40864 standard; protein; 237 AA.			
XX	AAM40864;			
XX	22-OCT-2001 (first entry)			
XX	Human polypeptide SEQ ID NO 5795.			
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;			
XX	peripheral nervous system; neuropathy; central nervous system; CNS;			
XX	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;			
XX	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;			
XX	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;			
XX	leukaemia.			
XX	Homo sapiens.			
OS	WO200153312-A1.			
XX	26-JUL-2001.			
XX	26-DEC-2000; 2000WO-US034263.			
XX	23-DEC-1999; 99US-00471275.			
PR	21-JAN-2000; 2000US-00488725.			
PR	25-APR-2000; 2000US-00552317.			
PR	20-JUN-2000; 2000US-00598042.			
PR	19-JUL-2000; 2000US-00620312.			
PR	03-AUG-2000; 2000US-00653450.			
PR	14-SEP-2000; 2000US-00662191.			
PR	19-OCT-2000; 2000US-00693036.			
PR	29-NOV-2000; 2000US-00727344.			
XX	(HYSE-) HYSEQ INC.			
XX	Wang Y, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;			
PI	Tang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;			
PI	Zhou P, Goodrich R, Drmanac RT;			
XX	WPI; 2001-442253/47.			
XX	N-PSDB; AAI60020.			
XX	Novel nucleic acids and polypeptides, useful for treating disorders such			
PPT	as central nervous system injuries.			
XX	Example 2; SEQ ID NO 5795; 10078pp; English.			
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the			
CC	encoded polypeptides (AAM38642-AAM42213) with nootropic,			
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful			
CC	in gene therapy. A composition containing a polypeptide or polynucleotide			
CC	of the invention may be used to treat diseases of the peripheral nervous			
CC	system, such as peripheral nervous injuries, peripheral neuropathy and			
CC	localised neuropathies and central nervous system diseases, such as			

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilization of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 XX Sequence 237 AA;

Query Match 68.2%; Score 1162; DB 4; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-97;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 TRSPVEIISAPHPHLLFLLKGYEAPQIALRCGIMLRCIRHEPLAKIILFSNPRDPFK 170  
 Db 2 TRSPVEIISAPHPHLLFLLKGYEAPQIALRCGIMLRCIRHEPLAKIILFSNPRDPFK 61  
 QY 171 YVELSTFDIASDAPATFKDILLTRHKVLVADFLQNYDTIFEDYEKLQSENVTKQSLK 230  
 Db 62 YVELSTFDIASDAPATFKDILLTRHKVLVADFLQNYDTIFEDYEKLQSENVTKQSLK 121  
 QY 231 LLGELILDRHPNIMTKYISKPENLKLMMNLLRDKSPNIQFEAPHVFKVVASPHKTOPI 290  
 Db 122 LLGELILDRHPNIMTKYISKPENLKLMMNLLRDKSPNIQFEAPHVFKVVASPHKTOPI 181  
 QY 291 VEILLKNPKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 337  
 Db 182 VEILLKNPKLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLKKTAP 228

## RESULT 12

ABB60392  
 ID ABB60392 standard; protein; 339 AA.

AC ABB60392;  
 XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 7968.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL04495.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 7968; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU160511), expressed DNA  
 CC sequences (ABU01840-ABU16175) and the encoded proteins (AB857737-  
 CC AB872072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 339 AA;

Query Match 65.2%; Score 1111; DB 4; Length 339;  
 Best Local Similarity 65.0%; Pred. No. 2.3e-92;  
 Matches 217; Conservative 59; Mismatches 54; Indels 4; Gaps 3;

QY 4 MPLESKSHKPAETVKILKONLALIEKQDKTKDYASSEVSKSLQAMKEILCGTNEKEPPT 63  
 Db 1 MPLEFGSKQKSPVEIUVKSLKEAINALEAGDRKVEKAQEDVSKNLVSIKNNLYGSSDABPPA 60  
 QY 64 E-AVAQAQLQYSGLELVTLIADQLIDFEGKDKDVTQIFNNILRROIGTSPTEVYISAH 122  
 Db 61 DYVVAQLSQEILYNSNLLLIQNLHRIDFEGKKHVALIFNNVLRQIGTSPTEVYICTK 120  
 QY 123 PHILFMLLKGYE--APOIALRCGIMLRECIHREPLAKIILFSNQFRDFFKYVELSTFDIA 180  
 Db 121 PEILFTLMAGYEDAHPEIATLNSGTMLRECIHREPLAKIILFSDEFFKFFRYVEVSTFDIA 180  
 QY 181 SDAPATEFKDILLTRHKVLVADFLQNYDTIF-EDYEKLQSENVTKQSLKLGELILDR 239  
 Db 181 SDAPSTFKELLTRHKVLCABFLDANYDKFFSQHYQRLNSENYYTRQSLKLGELILDR 240  
 QY 240 HNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAPHVFKVVASPHKTOPIVEILLKNQ 299  
 Db 241 HNFVTMTRYISEPENLKLMMNLLKESRNQFEAPHVFKVVASPHKTOPIVDILLRNQT 300  
 QY 300 KLIEFLSSFKERTDDEQFADEKKNYLIKQIRDLK 333  
 Db 301 KLVDLFTNFTDRSDEQFNDKAYLIKQIKELK 334

## RESULT 13

AA94249

ID AA94249 standard; protein; 339 AA.

AC AA94249;

DT 10-AUG-2000 (first entry)

DE Drosophila calcium binding protein DMO25.

XX Drosophila; calcium binding protein; cancer; inflammation; DMO25; CBP;  
 KW reproductive disorder; autoimmune disorder; developmental disorder;  
 KW seizure disorder; immune disorder; infection.

OS Drosophila melanogaster.

XX WO200029580-A1.

XX 25-MAY-2000.

PF 12-NOV-1999; 99WO-US027027.

XX 13-NOV-1998; 98US-00190965.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX WPI; 2000-387793/33.

XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
 PT diagnosis, prevention and treatment of cancers, immune, developmental or  
 PT reproductive disorders.

XX Disclosure; Page 67-68; 72pp; English.

XX The present sequence is the Drosophila calcium binding protein DMO25. It  
 CC was used in a sequence alignment to identify human calcium binding  
 CC protein hCBP. The hCBP protein and the gene encoding it are useful for  
 CC the diagnosis and treatment of the following types of disorder: cancers  
 CC (such as adenocarcinomas), reproductive disorders (such as infertility,  
 CC ovulatory defects, endometriosis, disruptions of the oestrus and  
 CC menstrual cycles), polycystic ovary syndrome and ovarian  
 CC hyperstimulation), autoimmune disorders (such as benign prostatic  
 CC hyperplasia and prostatitis), developmental disorders (such as Cushing's  
 CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary  
 CC neuropathies, seizure disorders, immune disorders (such as AIDS,  
 CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's  
 CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,  
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative  
 CC colitis), and viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections  
 XX  
 SQ Sequence 339 AA;  
 Query Match 65.1%; Score 1109; DB 3; Length 339;  
 Best Local Similarity 65.0%; Pred. No. 3.6e-92;  
 Matches 217; Conservative 59; Mismatches 54; Indels 4; Gaps 3;  
 QY 4 MPFSGKHNPATVVKILKDNLAILEKODKTKASEEVSKSLQAMKEITLQCTNEKEPPT 63  
 DB 1 MPFSGKSQSPVELVSKSLKEAENALEAGDKVKEQEDVSKNLVSKMLHSSDAEP 60  
 QY 64 E-AVLAQLAQLYSSGLLVTLADLQIDPEGKDVQTFIPNNLRQIGTRSPVVEISAH 122  
 DB 61 DYVVAQLSQELYNLNLILLIQNLHRIDPEGKHHVALIPNNLRQIGTRSPVVEICTK 120  
 QY 123 PHILFMLKGYE--APQALRCGIMRECIRHEPLAKILFNSQRPDFKYVELSTFDIA 180  
 DB 121 PEILFTLMAGYEDAHPEIALNSGTMLRECARVEALAKIMLHSDPEFKFRYVEVSTFDIA 180  
 QY 181 SDAFATFKDLTRHKVLVADFLQNVDTIF--EDYEKLQSENVTYKROSLKGLGELLDR 239  
 DB 181 SDAFSTFKELLTRHKLCAEFLDANYDKFFSHYQRLNSENVTYKROSLKGLGELLDR 240  
 QY 240 HNFATMTYISKPENKLMNNLRDKSPNIQFEAFHVKVVFVASHPKTQPIVEILLKNQ 299  
 DB 241 HNFVTWTRYISEPENKLMNNLRDKSRNIQFEAFHVKVVFVANENKPKPILDLILRNQ 300  
 QY 300 KLIEFLSSFOKERTDEQFADEKKNYLIQIRDLK 333  
 DB 301 KLVDFTLNFHTDRSDEQFNDEKAYLIQIKELK 334

RESULT 14  
 AAY94250  
 ID AAY94250 standard; protein; 377 AA.  
 XX  
 AC AAY94250;  
 XX  
 DT 10-AUG-2000 (first entry)  
 XX  
 DE C. elegans yeast-like calcium binding protein.  
 XX  
 KW Calcium binding protein; cancer; inflammation; yeast-like CBP; CBP;  
 KW reproductive disorder; autoimmune disorder; developmental disorder;  
 KW seizure disorder; immune disorder; infection.  
 OS Caenorhabditis elegans.  
 XX  
 PN WO200029580-A1.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PF 12-NOV-1999; 99WO-US027027.  
 XX  
 PR 13-NOV-1999; 98US-00190965.  
 XX

PA  
 XX  
 PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;  
 XX WPI; 2000-387793/33.  
 DR Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
 PT diagnosis, prevention and treatment of cancers, immune, developmental or  
 PT reproductive disorders.  
 PS Disclosure; Page 68-69; 72pp; English.  
 XX  
 CC The present sequence is the C. elegans yeast-like CBP. It was used in a  
 CC sequence alignment to identify human calcium binding protein hCBP. The  
 CC hCBP protein and the gene encoding it are useful for the diagnosis and  
 CC treatment of the following types of disorder: cancers (such as  
 CC adenocarcinomas), reproductive disorders (such as infertility, ovulatory  
 CC defects, endometriosis, disruptions of the oestrus and menstrual cycles,  
 CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune  
 CC disorders (such as benign prostatic hyperplasia and prostatitis),  
 CC developmental disorders (such as Cushing's syndrome, muscular dystrophy  
 CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,  
 CC immune disorders (such as AIDS, allergies, anaemia, asthma,  
 CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'  
 CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis,  
 CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,  
 CC bacterial, fungal, parasitic, protozoal and helminthic infections  
 XX  
 SQ Sequence 377 AA;  
 Query Match 62.4%; Score 1063.5; DB 3; Length 377;  
 Best Local Similarity 60.5%; Pred. No. 5.6e-88;  
 Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;  
 QY 4 MP-LFSKSHKNPATVVKILKDNLAILEK-----QDKTKDKASEEVSKSLQAM 49  
 DB 1 MPFLFGKSHKSPADVVKTREVLITLDPKPPKLDKQGNQSDKKYDKALDEVSKNVAMI 60  
 QY 50 KEILCGTNEKEPPTB---AVLAQLAQLYSSGLLVTLADLQIDPEGKDVQTFIPNNLR 106  
 DB 61 KSFTYGNDSAPSSSHVQVLAQAEVYVNAVILPMLIRPKLPECKDKVQGIENLLR 120  
 QY 107 ROIGTRSPVVEISAHPHILFMLKGYEAPQALRCGIMRECIRHEPLAKILFNSQ 166  
 DB 121 ROIGTRSPVVEISAHPHILFMLKGYEAPQALRCGIMRECIRHEPLAKILFNSQ 180  
 QY 167 DFFKYVELSTFDIASDAPATFKDLTRHKVLVADFLQNVDTIFEDYEKLQSENVTYK 226  
 DB 181 TFFLYVQSEVDFISSDAPSTFKELTTRHKAIIAEFLDSNYDTFFAQYQNLNLSKNVYTR 240  
 QY 227 QSLKGLGELLILDRHNFATMTYISKPENKLMNNLRDKSPNIQFEAFHVKVVFVASHPK 286  
 DB 241 QSLKGLGELLILDRHNFATMTYISKPENKLMNNLRDKSRNIQFEAFHVKVVFVANENK 300  
 QY 287 TQPIVEILLKNQPKLIEFLSSFOKERTDEQFADEKKNYLIQIRDLKKT 335  
 DB 301 KPISDILNRNREKLVEFLSEFHNDRITDDEQFNDEKAYLIQIQEMKSS 349

RESULT 15  
 AAG45273  
 ID AAG45273 standard; protein; 343 AA.  
 XX  
 AC AAG45273;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 56816.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX

OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121822P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 07-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.

Job time : 63 secs

```

PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161408P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.0%; Score 716.5; DB 3; Length 343;
Best Local Similarity 42.9%; Pred. No. 1.7e-56;
Matches 144; Conservative 78; Mismatches 105; Indels 9; Gaps 3;

Qy 6 LPSKSHKPAEIVKILKNLAILEK-----QDKKTKASEEVSKSLQAMKEILCGTNE 58
Db 4 LFKSKPRTADIVRQTRDLLLYADRSNFPDLRESKREKVKVLSKSIDLKLILYGNSE 63

Qy 59 KEPTEAVAQALQELYSGLAVTLADIQLIDFGKKDVTQIFNNILRROIGTRSPVVEY 118
Db 64 AEPVAEACQATQEFFKADTLRLLTSLPNNLEARKDQVAVANLQKQVNSRLIADY 123

Qy 119 ISAPHILFMLLKGVEAFQIALRCIMLRECIHREPLAKIILFSNQPRDFKVELSTFD 178
Db 124 LESNIDLMDFVDCGFENTDMALHYGTMFECIRHQIVAKYVLDSEHVKKFFYYIQLPNFD 183

Qy 179 IASDAFATFKOLLTRHKVULVADFLBQNYDTTFEDY-EKLOSENVYTRQSLKLGELIL 237
Db 184 IADAAATFKELLTRHKSTVAEFLTKNEDWFFADYNSKLESTNYITRRQAIKLLGDILL 243

Qy 238 DRHNFAMTKYISKPENIKLMMNLRLDKSPNIOEAFHVKVFEVASPHKTOPIVEILLKN 297
Db 244 DRNSAVMTKVSSMDNIRILNLLRESSTKIQIEAFHVKLFVANQNKPSDIANILVAN 303

Qy 298 QPKLIEFLSSFKERTDDEQFADEKNYLIKQIRDLK 333
Db 304 RNKLLRLADIKPKD-EDERFDADKAQVREIANLK 338

```

Search completed: April 12, 2004, 10:26:39



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:27:48 ; Search time 46 Seconds  
(without alignments)  
1926.306 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 1704

Sequence: 1 MKKMPLFSKSHKNPAEIVKI.....FADEKNYLIKQIDLLKKTAP 337

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1073127 seqs, 262937947 residues

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	100.0	337	14	US-10-025-730-1
2	1381	81.0	496	14	Sequence 1, Appli
3	1381	81.0	552	14	Sequence 5, Appli
4	1376	80.8	341	14	Sequence 6, Appli
5	1109	65.1	339	14	Sequence 3, Appli
6	1063.5	62.4	377	14	Sequence 4, Appli
7	756	44.4	412	12	Sequence 5, Appli
8	752	44.1	446	12	US-10-025-730-5
9	715.5	42.0	339	12	Sequence 54669, A
10	649	38.1	351	12	Sequence 52177, A
11	630.5	37.0	342	12	Sequence 224185, A
12	514	30.2	337	12	Sequence 46025, A
13	395.5	23.2	173	12	Sequence 48972, A
14	322.5	18.9	184	12	Sequence 209662, A
15	275.5	16.2	161	12	Sequence 158029, A
					Sequence 149595, A
					Sequence 70029, A

16	227.5	13.4	135	11	US-09-864-408A-6108
17	186.5	10.9	83	12	US-10-425-114-44450
18	153	9.0	74	12	US-10-425-114-56749
19	131.5	7.7	63	12	US-10-424-599-230655
20	128.5	7.5	820	14	US-10-029-386-32324
21	128.5	7.5	3878	14	US-10-080-608A-11
22	128.5	7.5	3899	14	US-10-171-311-4
23	128.5	7.5	3907	14	US-10-171-311-2
24	128.5	7.5	3911	15	US-10-370-685-100
25	128.5	7.5	3917	14	US-10-171-311-8
26	128.5	7.5	3925	14	US-10-171-311-6
27	126	7.4	116	12	US-10-424-599-210348
28	125	7.3	660	9	US-09-864-761-47959
29	120.5	7.1	1182	12	US-10-282-122A-53445
30	116.5	6.8	709	15	US-10-023-634-18
31	116.5	6.8	709	15	US-10-023-634-77
32	116.5	6.8	724	15	US-10-023-634-80
33	113.5	6.7	57	12	US-10-424-599-224186
34	113	6.6	2184	14	US-10-304-095-6
35	111.5	6.5	725	15	US-10-023-634-79
36	111.5	6.5	725	15	US-10-205-647A-4
37	108	6.3	769	14	US-10-032-585-7117
38	107	6.3	980	12	US-10-211-462-145
39	107	6.3	1128	12	US-10-282-122A-69138
40	106.5	6.2	709	14	US-10-256-250-15
41	106.5	6.2	709	15	US-10-023-634-14
42	106.5	6.2	725	10	US-09-978-309A-47
43	106.5	6.2	725	12	US-10-257-377-1
44	106.5	6.2	725	14	US-10-256-250-14
45	106.5	6.2	725	15	US-10-023-634-78

#### ALIGNMENTS

#### RESULT 1

US-10-025-730-1  
; Sequence 1, Application US/10025730  
; Publication NO. US20030045466A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/10/025,730  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: US/09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 3734805  
US-10-025-730-1

Query Match 100.0%; Score 1704; DB 14; Length 337;  
Best Local Similarity 100.0%; Pred. No. 5.5e-147;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKKMPLFSKSHKNPAEIVKILKDNLAIEKQDKTKDASEVSKSLQAMKEILCTGNEKE	60
Db	1	MKKMPLFSKSHKNPAEIVKILKDNLAIEKQDKTKDASEVSKSLQAMKEILCTGNEKE	60
Qy	61	PPTAEVAQAQELYSGLLVTLIADQLIDFEGKDVQTFIENILRRQIGTRSPVYEIS	120
Db	61	PPTAEVAQAQELYSGLLVTLIADQLIDFEGKDVQTFIENILRRQIGTRSPVYEIS	120
Qy	121	AHPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFSNQPRDFKVELSTFDIA	180

Db 121 AAPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQNRDPFKYVELSTFDIA 180  
QY 181 SDAPATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDRH 240  
Db 181 SDAPATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDRH 240  
QY 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 300  
Db 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 300  
QY 301 LIEFLSSFOKERTDDEQFADEKNYLILKQIRDLLKKTAP 337  
Db 301 LIEFLSSFOKERTDDEQFADEKNYLILKQIRDLLKKTAP 337

## RESULT 2

US-10-239-079-5  
; Sequence 5, Application US/10239079  
; Publication No. US20030148446A1  
; GENERAL INFORMATION:  
; APPLICANT: Merck Patent GmbH  
; TITLE OF INVENTION: ANIC-BP1-ligand  
; FILE REFERENCE: ANIC-BP-1-ligand  
; CURRENT APPLICATION NUMBER: US/10/239,079  
; CURRENT FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Gal4-ANIC-BP-1  
US-10-239-079-5

Query Match 81.0%; Score 1381; DB 14; Length 496;  
Best Local Similarity 81.0%; Pred. No. 2.8e-117;  
Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;  
QY 4 MPL-PSKSHKPAEIVKILKDNLAILEKQ---DKTKDASEEVSLSQAMKEILCGTNEK 59  
Db 156 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKAEKATEEVSQNLVAMKEILYGTNEK 215  
QY 60 EPPTAEVAQLAQELYSGLLVTLIADQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEYI 119  
Db 216 EPQTEAVAQLAQELYSGLLVTLIADQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEYI 275  
QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQNRDPFKYVELSTFDI 179  
Db 276 CTQQNILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQNRDPFKYVELSTFDI 335  
QY 180 ASDAFATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDR 239  
Db 336 ASDAFATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDR 395  
QY 240 HNFAMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 299  
Db 396 HNFAMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 455  
QY 300 LIEFLSSFOKERTDDEQFADEKNYLILKQIRDLLKKTAP 336  
Db 456 KLIEFLSKFQNDRTDEQFDEKNTYLVKQIRDLKRP 492

## RESULT 3

US-10-239-079-6  
; Sequence 6, Application US/10239079  
; Publication No. US20030148446A1  
; GENERAL INFORMATION:  
; APPLICANT: Merck Patent GmbH  
; TITLE OF INVENTION: ANIC-BP1-ligand

; FILE REFERENCE: ANIC-BP-1-ligand  
; CURRENT APPLICATION NUMBER: US/10/239,079  
; CURRENT FILING DATE: 2002-09-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: LexA-ANIC-BP-1  
US-10-239-079-6

Query Match 81.0%; Score 1381; DB 14; Length 552;  
Best Local Similarity 81.0%; Pred. No. 3.3e-117;  
Matches 273; Conservative 31; Mismatches 29; Indels 4; Gaps 2;  
QY 4 MPL-PSKSHKPAEIVKILKDNLAILEKQ---DKTKDASEEVSLSQAMKEILCGTNEK 59  
Db 212 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKAEKATEEVSQNLVAMKEILYGTNEK 271  
QY 60 EPPTAEVAQLAQELYSGLLVTLIADQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEYI 119  
Db 272 EPQTEAVAQLAQELYSGLLVTLIADQLIDFEGKQDVTOIFNNILRRQIGTRSPVVEYI 331  
QY 120 SAHPHILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQNRDPFKYVELSTFDI 179  
Db 332 CTQQNILFMLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSQNRDPFKYVELSTFDI 391  
QY 180 ASDAFATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDR 239  
Db 392 ASDAFATFKDLLTRHKVYLVADFLQNYDTIFEDYEKLQSENVTYTKRQSLKLGELILDR 451  
QY 240 HNFAMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 299  
Db 452 HNFAMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKFVVASPHKTPQIVIEILLKNQPK 511  
QY 300 KLIEFLSSFOKERTDDEQFADEKNYLILKQIRDLLKKTAP 336  
Db 512 KLIEFLSKFQNDRTDEQFDEKNTYLVKQIRDLKRP 548

## RESULT 4

US-10-025-730-3  
; Sequence 3, Application US/10025730  
; Publication No. US20030045466A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/10/025,730  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: US/09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; FEATURE:  
; OTHER INFORMATION: g262934  
US-10-025-730-3

Query Match 80.8%; Score 1376; DB 14; Length 341;

Best Local Similarity 80.7%; Pred. No. 4.7e-117;  
Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;  
QY 4 MPL-PSKSHKPAEIVKILKDNLAILEKQ---DKTKDASEEVSLSQAMKEILCGTNEK 59

Db 1 MPFFGKSHKSPADIVKVKESMAVLEKQSDKAEKATEEVSXNVLAMKEILYGTNEK 60  
Qy 60 BPTEAVAQLAQELYSGLLVTLIADLQIDPEGKQVTOIFNNILRRQIGTRSPTEYI 119  
Db 61 EPQTEAVAQLAQELYSGLLVTLIADLQIDPEGKQVTOIFNNILRRQIGTRSPTEYI 120  
Qy 120 SAHPHILPMLKGYEAPQIALRCGIMLRECIHREPLAKIILFSNQFRDFPKYVELSTPDI 179  
Db 121 CTOQNILPMLKGYEAPQIALRCGIMLRECIHREPLAKIILFSNQFRDFPKYVELSTPDI 180  
Qy 180 ASDAFATKOLLTRHKVLVADFLQNYDTIFEDYKLLQSENYYTKRQSLKLLGELLIDR 239  
Db 181 ASDAFATKOLLTRHKVLVADFLQNYDTIFEDYKLLQSENYYTKRQSLKLLGELLIDR 240  
Qy 240 HNFALMTKYISKPENLKLMMNLLRDKSPNIOPEAFHVKFVVASPHKTOPIVEILLKNQ 299  
Db 241 HNFALMTKYISKPENLKLMMNLLRDKSRNIOPEAFHVKFVVASPHKTOPIVEILLKNQ 300  
Qy 300 KLIFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 336  
Db 301 KLIFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 337

## RESULT 5

US-10-025-730-4

; Sequence 4, Application US/10025730  
; Publication No. US20030045466A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 4

; LENGTH: 339

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE: -

; OTHER INFORMATION: g1794137

US-10-025-730-4

Query Match 65.1%; Score 1109; DB 14; Length 339;  
Best Local Similarity 65.0%; Pred. No. 1.1e-92;  
Matches 217; Conservative 59; Mismatches 54; Indels 4; Gaps 3;

Qy 4 MPFESKSHKPAEIVKILKDNLAILEKQDKKTDKASEEVSXSLQAMKBILOCTNEKEPPT 63  
Db 1 MPFESKSHKPAEIVKILKDNLAILEKQDKKTDKASEEVSXSLQAMKBILOCTNEKEPPT 60  
Qy 64 E-AVAQLAQELYSGLLVTLIADLQIDPEGKQVTOIFNNILRRQIGTRSPTEYI 122  
Db 61 DYVVAQLSQELYSGLLVTLIADLQIDPEGKQVTOIFNNILRRQIGTRSPTEYI 120  
Qy 123 PHILPMLKGYE--APQIALRCGIMLRECIHREPLAKIILFSNQFRDFPKYVELSTPDI 180  
Db 121 PHILPMLKGYE--APQIALRCGIMLRECIHREPLAKIILFSNQFRDFPKYVELSTPDI 180  
Qy 181 SDAFATKOLLTRHKVLVADFLQNYDTIF--EDYKLLQSENYYTKRQSLKLLGELLIDR 239  
Db 181 SDAFATKOLLTRHKVLVADFLQNYDTIF--EDYKLLQSENYYTKRQSLKLLGELLIDR 240  
Qy 240 HNFALMTKYISKPENLKLMMNLLRDKSPNIOPEAFHVKFVVASPHKTOPIVEILLKNQ 299  
Db 241 HNFALMTKYISKPENLKLMMNLLRDKSRNIOPEAFHVKFVVASPHKTOPIVEILLKNQ 300

Qy 300 KLIFLSSFOKERTDDEQFADEKKNYLIKQIRDLK 333  
Db 301 KLVDFLTNFHTDRSEDEQFNDEKAYLIKQIKELK 334

## RESULT 6

US-10-025-730-5

; Sequence 5, Application US/10025730

; Publication No. US20030045466A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 5

; LENGTH: 377

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

; FEATURE: -

; OTHER INFORMATION: g1255838

US-10-025-730-5

Query Match 62.4%; Score 1063.5; DB 14; Length 377;

Best Local Similarity 60.5%; Pred. No. 1.8e-88;

Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;

Qy 4 MP-LFESKSHKPAEIVKILKDNLAILEK-----QDKKTDKASEEVSXSLQAM 49  
Db 1 MP-LFESKSHKPAEIVKILKDNLAILEK-----QDKKTDKASEEVSXSLQAM 60  
Qy 50 KEILCGTNEKEPPTPE---AVAQLAQELYSGLLVTLIADLQIDPEGKQVTOIFNNILR 106  
Db 61 KSFTYGNDSAEPSSEHVVQVAQLAQEVNANILPMLKLPKFEFECKQVGFNNLLR 120  
Qy 107 ROIGTRSPTEYI SAHPHILPMLKGYEAPQIALRCGIMLRECIHREPLAKIILFSNQFR 166  
Db 121 ROIGTRSPTEYI LGARPEILQLVQGVSPDIALTCGLMLRESIRHDLAKIILYSDFY 180  
Qy 167 DFYKVELSTPDIASDAFATKOLLTRHKVLVADFLQNYDTIFEDYKLLQSENYYTKR 226  
Db 181 TFELYVQSEVFDISSDAFATKOLLTRHKVLVADFLQNYDTIFEDYKLLQSENYYTKR 240  
Qy 227 QSLKLLGELLIDRHNFAIMTKYISKPENLKLMMNLLRDKSPNIOPEAFHVKFVVASPHK 286  
Db 241 QSLKLLGELLIDRHNFAIMTKYISKPENLKLMMNLLRDKSRNIOPEAFHVKFVVASPHK 300  
Qy 287 TOPIVEILLKNQPKLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 335  
Db 301 KPISDILNRNREKLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKT 349

## RESULT 7

US-10-425-114-54669

; Sequence 54669, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54669
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3069-003-D7_FLI.pap
US-10-425-114-54669

Query Match
  44.4%; Score 756; DB 12; Length 412;
Best Local Similarity 45.3%; Pred. No. 2.3e-60;
Matches 153; Conservative 67; Mismatches 110; Indels 8; Gaps 3;

QY 2 KMPLEFSKSHQPAEIVKILKDNLAILE-----KQDKTKDASEVSKSLQAMKEILCG 55
Db 76 KMKGFLFKTPRTVDIVRQTRCLVHLDLHSGSRSGDAKDEKMTLSKNIRDMKSILYG 135
QY 56 TNEKEPPTAEVAQAOLAEYSSGLLVTLIADLQIDFEGKKDVTQIFNNILRQIGTRSP 115
Db 136 NGESEPVTEACVQLTQEFFRENTLRLLIHLPKUNLETRKDQATQVANLQORQVSSKIVA 195
QY 116 VEYISAHPHILFMLKGYEAPQIALRCGIMLRECIHREPLAKIILFNSQFRDFFKYVELS 175
Db 196 SEYLESNKDLLDILGVENMDIALHYGAMLRERHQSIRYVLESEHMKKFFDYIQLP 255
QY 176 TFDIASDAFATFKDLLTRHKVLVADFLQNYDTIFEDY-EKLLQSENVYTKRQSLKLGE 234
Db 256 NFDIASDASATFKELLTRHKATVAEFLSNNDYDWFEEFNSRLLSSTNYITKQAIKLLGD 315
QY 235 LILDRHFAINTKYISKPENILKMNLLRDSPNIQFEAFHVKVVFASPHKTQPIVEIL 294
Db 316 MLLDRSNVAVMRYVSSKDNLMILNLRDSSKNQIIEAFHVKLFPAANKNKPPEVNIL 375
QY 295 LKNQPKLIEFLSSFOKERTDDEQFADEKNYLIKQIRDL 332
Db 376 VTRNKKLRFPAFGKIDK-EDEQFEADKEHVIKEISAL 412

RESULT 8
US-10-425-114-52177
; Sequence 52177, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52177
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-007-H10_FLI.pap
US-10-425-114-52177

Query Match
  44.1%; Score 752; DB 12; Length 446;
Best Local Similarity 45.0%; Pred. No. 5.9e-60;
Matches 152; Conservative 67; Mismatches 111; Indels 8; Gaps 3;

QY 2 KMPLEFSKSHQPAEIVKILKDNLAILE-----KQDKTKDASEVSKSLQAMKEILCG 55
Db 110 KMKGFLFKTPRTVDIVRQTRCLVHLDLHSGSRSGDAKDEKMTLSKNIRDMKSILYG 169

```

```

QY 56 TNEKEPPTAEVAQAOLAEYSSGLLVTLIADLQIDFEGKKDVTQIFNNILRQIGTRSP 115
Db 170 NGESEPVTEACVQLTQEFFRENTLRLLIHLPKUNLETRKDQATQVANLQORQVSSKIVA 229
QY 116 VEYISAHPHILFMLKGYEAPQIALRCGIMLRECIHREPLAKIILFNSQFRDFFKYVELS 175
Db 230 SEYLESNKDLLDILGVENMDIALHYGAMLRERHQSIRYVLESEHMKKFFDYIQLP 289
QY 176 TFDIASDAFATFKDLLTRHKVLVADFLQNYDTIFEDY-EKLLQSENVYTKRQSLKLGE 234
Db 290 NFDIASDASATFKELLTRHKATVAEFLSNNDYDWFEEFNSRLLSSTNYITKQAIKLLGD 349
QY 235 LILDRHFAINTKYISKPENILKMNLLRDSPNIQFEAFHVKVVFASPHKTQPIVEIL 294
Db 350 MLLDRSNVAVMRYVSSKDNLMILNLRDSSKNQIIEAFHVKLFPAANKNKPPEVNIL 409
QY 295 LKNQPKLIEFLSSFOKERTDDEQFADEKNYLIKQIRDL 332
Db 410 VTRNKKLRFPAFGKIDK-EDEQFEADKEHVIKEISAL 446

RESULT 9
US-10-424-599-224185
; Sequence 224185, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224185
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_44468C.1.pap
US-10-424-599-224185

Query Match
  42.0%; Score 715.5; DB 12; Length 339;
Best Local Similarity 46.3%; Pred. No. 8.5e-57;
Matches 151; Conservative 61; Mismatches 109; Indels 5; Gaps 2;

QY 12 KNPAEIVKILKDNLAILEKQD----KKTDKASEVSKSLQAMKEILCOTNEKEPTEAVA 67
Db 12 KTPQEVAKSIKESLMALDTKTVVEVKALEEVEKNFVTRMTLSGDSEPNLDQVS 71
QY 68 QLAQELVSSGLLVTLIADLQIDFEGKKDVTQIFNNILRQIGTRSPTEVYISAHPHILF 127
Db 72 QLVBEICKEDVTLTIHLKLPILGWARKDLVHCKVETNYCYVEIQHIELLD 131
QY 128 MLLKGYEAPQIALRCGIMLRECIHREPLAKIILFNSQFRDFFKYVELSTPDASDAPATF 187
Db 132 FLVVCYDNKDIALSCGIMLRECIKFPPLARILSASFVLFKVELPNFVDADFSTF 191
QY 188 KDLTRHKVLVADFLQNYDTIFEDYKLLQSENVYTKRQSLKLGEILDRHNFAMTK 247
Db 192 KDLTKHVNVSSEFLTAHYDEFFDLYEKLLTSPNVVTRQSLKLLSEFLSPNSQIMQ 251
QY 248 YISKPENILKMNLLRDSPNIQFEAFHVKVVFASPHKTQPIVEILLKQKLEFLSS 307
Db 252 YILEVRYLKVMVMTLLRDSSKNQIISAPHIFKVFVANPNKPREVKIILSKNOEKLDDLHN 311
QY 308 FOKER-TDDEQFADEKNYLIKQIRDL 332
Db 312 LSPGGSDEQFEKEEFLIKEIERL 337

```

```
RESULT 10
US-10-425-114-46025
; Sequence 46025, Application US/10425114
; Publication No. US20040034868A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46025
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700170954_FLI.pep
US-10-425-114-46025

Query Match      38.1%; Score 649; DB 12; Length 351;
Best Local Similarity 42.3%; Pred. No. 1e-50;
Matches 138; Conservative 67; Mismatches 117; Indels 4; Gaps 2;

QY      8 SKSHKNPAIIVKILNDLAILEKQDKTD-KASEEVSKSLOAMKEILCGTNEKEPTEAV 66
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      24 SRQRLPQEIARSLSKLSVAL--DTKTGAKALEDAEAKNLTIRHTLAGDGEVPEPQEV 80
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      67 AQLAELYSSGLLVTLADLQIDPEGKDVQTFNNILRROIGTRSPVTEYISAPHTL 126
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      81 LQALAEICKEGVLSPVQMLPSLGMWARKDLVHCWICILLRQKQVDESYCCVQVYENHFDLL 140
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      127 FMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFSNQPRDFKVKYVELSTFDIASDAFAT 186
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      141 DFLVVCYKNLEVALNGMRLCEIKYPTLAKYLESSSEFELFQYVELSNFSDALNT 200
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      187 FKDLTRHKVLVADFLQNYDTIFEDYEKLGLOSENVYTKRQSLKLGELILDRHNFAMT 246
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      201 FKDLTLKHEDAVSEFLSSHYEQFFGLYTKLSSSTNVYTRRQSKVSEFLLEAPNAQIMK 260
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      247 KYISKPENLKLMMNLDRKSPNIQPEAFHVKFVASFHKTQPIVEILLKQPKLIEFLS 306
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      261 RYILEVHYLNIMWGLLKSSKNIRICAFHFVKFVANPNKPREIIQFLVENHREVLKLLH 320
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      307 SPQKERTDDEQFADEKNYLIKQIRD 332
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      321 NLPTSKEGEQDLDEERDLIIEKEL 346
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 11
US-10-425-114-48872
; Sequence 48872, Application US/10425114
; Publication No. US20040034868A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48872
; LENGTH: 342

Query Match      30.2%; Score 514; DB 12; Length 337;
Best Local Similarity 35.4%; Pred. No. 2e-38;
Matches 120; Conservative 79; Mismatches 122; Indels 18; Gaps 10;

QY      6 LFSKSHKNPAETIVKILNDLAILE-----KQDKTKDKASEEVSKSLOAMKEILCGTNEKE 60
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      5 LFSKSKERTPPDVVRQTRDLLLFVDRSPETRESKREKQWSELFKNIRELKSILYGNSSSE 64
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      61 PPTTEAVA-QLAQELVSSGLLVTLADLQIDPEFGKQVQTFNNILRROIGTRSPTEYI 119
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      65 FVLRXCRQLTQTFKEDTLRLLIKLPKLNLEARKDATQVVAN-LRKQHSIMLSDYX 123
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

```
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3637-244-El_FLI.pep
US-10-425-114-48872

Query Match      37.0%; Score 630.5; DB 12; Length 342;
Best Local Similarity 42.0%; Pred. No. 4.9e-49;
Matches 134; Conservative 69; Mismatches 113; Indels 3; Gaps 2;

QY      14 PAEIVKILNDLAILEKQDKTDKASEEVSKSLOAMKEILCGTNEKEPTEAVAQLAQEL 73
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      26 PEEVRSIKSDSPALHTRTHA--KALEVEKNMSSRLILFGDGEVPENEEQVLIQITL 83
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      74 YSSGLLVTLADLQIDPEGKDVQTFNNILRROIGTRSPTEYISAPHTLFLMLKCY 133
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      84 CKEDVISLIVQDPLSLGWSGRKDLAICWICILRQKVDETYCCVQYLENHLLELDFLVGCY 143
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      134 EAPQIALRCGIMLRCEIRHEPLAKIILFSNQPRDFKVKYVELSTFDIASDAFATFKDILLR 193
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      144 KNLDIALNGMRLCEIKYPTLAKYLESGSELEFEVVELFNFDIASDALNTFKDILLTK 203
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      194 HKVLVADFLQNYDTIFEDYEKLGLOSENVYTKRQSLKLGELILDRHNFAMTKYISKE 253
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      204 HEVVVAEFLSSHYEQFFELYSLSSSTNVYTRRQAKFLSEFLLETHNSQIMKRVIVEVR 263
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      254 NLKLMNLLDRKSPNIQPEAFHVKFVASFHKTQPIVEILLKQPKLIEFLSSSQKERT 313
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      264 FLNIMINLLKSSKNIRICAFHFVKFVANPNKPRECIIVALLDNRRREVLLKLNLFSSKG 323
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      314 DDEQFADEKNYLIKQIRD 332
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      324 DDE-LDEEKDLIIQIQKL 341
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 12
US-10-424-599-209662
; Sequence 209662, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 209662
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (337)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_31352C.1.pep
US-10-424-599-209662

Query Match      30.2%; Score 514; DB 12; Length 337;
Best Local Similarity 35.4%; Pred. No. 2e-38;
Matches 120; Conservative 79; Mismatches 122; Indels 18; Gaps 10;

QY      6 LFSKSHKNPAETIVKILNDLAILE-----KQDKTKDKASEEVSKSLOAMKEILCGTNEKE 60
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      5 LFSKSKERTPPDVVRQTRDLLLFVDRSPETRESKREKQWSELFKNIRELKSILYGNSSSE 64
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      61 PPTTEAVA-QLAQELVSSGLLVTLADLQIDPEFGKQVQTFNNILRROIGTRSPTEYI 119
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB      65 FVLRXCRQLTQTFKEDTLRLLIKLPKLNLEARKDATQVVAN-LRKQHSIMLSDYX 123
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

```
QY 120 SAHPHLEWLLKGYEAPQ--TALRCG-IWLRECIHHEPLAKILLPS--NQFRD-PFKYVEL 174
DB 124 RVNWMFMBFLVVDYDEPHVILIRGAVLVGNASEHQIVSKYSSXLPKFLDYFFHPHQH 183
QY 175 STFDIASAFATFKDILLTRHKVLVADFLQNYDTTFEDYE-KLQSENKVTYKROSLKLLG 233
DB 184 PTFDTAANVTA-----ILTFHKSTEAFLSENWYFAEYNTKLESSNYITRQAVKLLG 239
QY 234 ELILDRHFAIMTKYISKPENIKLMMNLRLDRKSPNIOFEAFHVKFVVASPHKTOPIVEI 293
DB 240 DMLDRSNSAVMTRVYSSRDNRIILMNLIRRESSKSIQIEAFHVKFLFAANQHKPADIVSI 299
QY 294 LLKNOPKLEIFLSSFKERTDDEQFADEKKNYLKQIRD 332
DB 300 FVANKSKMLRLLEDFKIDK-EDEQFEADKAQVMREIAL 337

RESULT 13
US-10-424-599-158029
; Sequence 158029, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158029
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_11371C.1.pap
US-10-424-599-158029

Query Match 23.2%; Score 395.5; DB 12; Length 173;
Best Local Similarity 50.0%; Pred. No. 5.1e-28;
Matches 79; Conservative 35; Mismatches 37; Indels 7; Gaps 3;

QY 181 SDAPATFKDILLTRHKVLVADFLQNYDTTFEDYE-KLQSENKVTYKROSLKLLGILDR 239
DB 16 SDAATFKELLTRKSTVAEFLSKNWEFFAEYNTKLESSNYITRQAVKLLGMLDR 75
QY 240 HNPATMTYISKPENIKLMMNLRLDRKSPNIOFEAFHVKFVVASPHKTOPIVEILLKNOP 299
DB 76 SNSAVMTRVYSSRDNRIILMNLIRRESSKSIQIEAFHVKFLFAANQHKPADIVSI 135
QY 300 KLIEFLSSFKERTDDEQFADEKKNYLKQI-----RDL 332
DB 136 KLLRLGLDKLTKD-EDEQFEADKAQVIKEIAALEPRDL 172

RESULT 14
US-10-424-599-149595
; Sequence 149595, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```

```
; SEQ ID NO 149595
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106106C.1.pap
US-10-424-599-149595

Query Match 18.9%; Score 322.5; DB 12; Length 184;
Best Local Similarity 38.8%; Pred. No. 2.6e-21;
Matches 71; Conservative 37; Mismatches 72; Indels 3; Gaps 1;

QY 6 LFSKSHKNPABIVKILKNLAILEKQDKKTDKASEEVSKSQAAMEILCGTNEKEPPTPEA 65
DB 4 LFKPKRTPSDIVKQTRDILLRLTSRD---DDNMPDLTKNLRDLKSLYGNSESEFPVEA 60
QY 66 VAQLAQELYSGLLVTLIADLQLIDFEGKQDVOTQFNILRRQIGTRSPVTEYISAPHI 125
DB 61 CAQLTQEFADNTLRLLIQYLPKLNLEARKDATQIVANLQROVQSKLIASDYLDTNLDL 120
QY 126 LFMLLKGYEAPQIALRCGIMLRGIRHEPLAKIILFNSQFRDFPKYVELSTEDIASDAFA 185
DB 121 MDVLVSGYENTDMALHYGAMLRGIRHQIVAKYVLDSPHMKKVFYIQLPNFDIAADAAA 180
QY 186 TFK 188
DB 181 LLK 183

RESULT 15
US-10-425-114-70029
; Sequence 70029, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70029
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73031E03_FLI.pap
US-10-425-114-70029

Query Match 16.2%; Score 275.5; DB 12; Length 161;
Best Local Similarity 38.1%; Pred. No. 4.1e-17;
Matches 67; Conservative 34; Mismatches 54; Indels 21; Gaps 5;

QY 158 IILFNSQFRDFPKYVELSTEDIASDAFATFKDILLTRHKVLVADFLQNYDTTF-EDYEKL 216
DB 6 IILFYRMDSCFTLELL-----LNTNAYFFHTDL-----LCTDVVEQREKIFYQDSHYM 55
QY 217 LQSENYVTYKROSLKLLGILDRHNFATMTKYISKPENIKLMMNLRLDRKSPNIOFEAFH 276
DB 56 I-----FSFQLGDMLLDRSNAVMTRVSSKDNLMILMNLRLDRSSKNIOFEAFH 106
QY 277 FKVVFASPHKTOPIVEILLKNQPKLIEFLSSFKERTDDEQFADEKKNYLKQIRD 332
DB 107 FKLFAANKKPEVVNILLVTRNSKLLRFFAGFKIDK-EDEQFEADKEQVKEISAL 161

Search completed: April 12, 2004, 10:33:56
Job time : 47 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:27:13 ; Search time 22 Seconds  
(without alignments)  
790.816 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 1704

Sequence: 1 MKKMPFLFSKHNPAIVK.....FADEKNYLIKQIRDLKKTAP 337

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/prodata/2/iaa/PCRB COMB.pap.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	100.0	337	3	US-09-190-965-1
2	1704	100.0	337	4	US-09-470-253-1
3	1376	80.8	341	3	US-09-190-965-3
4	1376	80.8	341	4	US-09-470-253-3
5	1109	65.1	339	3	US-09-190-965-4
6	1109	65.1	339	4	US-09-470-253-4
7	1063.5	62.4	377	3	US-09-190-965-5
8	1063.5	62.4	377	4	US-09-470-253-5
9	128.5	7.5	3878	4	US-09-914-259-11
10	113.5	6.7	1279	4	US-09-724-517-2
11	113.5	6.7	1279	4	US-09-641-807A-2
12	113.5	6.7	1279	4	US-09-723-096-2
13	113	6.6	2184	4	US-09-417-485D-6
14	107	6.3	1155	4	US-09-543-681A-6286
15	105	6.2	586	2	US-08-630-822A-70
16	105	6.2	586	2	US-09-005-069-70
17	105	6.2	586	4	US-09-471-156A-30
18	105	6.2	586	4	US-09-004-730A-30
19	103.5	6.1	245	4	US-08-981-799A-30
20	103.5	6.1	245	4	US-09-399-913-4
21	103.5	6.1	245	4	US-09-298-731-4
22	103	6.0	387	4	US-09-328-352-5367
23	103	6.0	2682	4	US-09-595-684B-31
24	102.5	6.0	975	4	US-09-914-259-19
25	102.5	6.0	1098	3	US-08-923-992A-8
26	102.5	6.0	1164	3	US-08-923-992A-10
27	102.5	6.0	1388	4	US-09-572-191-2

```

28 102.5 6.0 1388 4 US-09-723-262-2 Sequence 2, Appli
29 102.5 6.0 1388 4 US-09-723-219-2 Sequence 2, Appli
30 102 6.0 474 3 US-08-387-117-6 Sequence 6, Appli
31 102 6.0 1128 3 US-08-923-992A-6 Sequence 6, Appli
32 101.5 6.0 1298 4 US-09-540-236-2334 Sequence 2334, Ap
33 101.5 6.0 1972 4 US-08-875-435B-4 Sequence 4, Appli
34 101 5.9 1147 3 US-08-470-260-5 Sequence 5, Appli
35 101 5.9 1147 3 US-08-471-491-5 Sequence 5, Appli
36 101 5.9 1147 3 US-08-466-662-5 Sequence 5, Appli
37 101 5.9 3289 2 US-08-477-451-2 Sequence 2, Appli
38 99.5 5.8 1164 3 US-08-923-992A-2 Sequence 2, Appli
39 99 5.8 323 4 US-09-134-001C-3133 Sequence 3133, Ap
40 99 5.8 728 4 US-09-711-164-300 Sequence 300, App
41 98 5.8 1048 3 US-09-356-952-5 Sequence 5, Appli
42 97.5 5.7 496 4 US-09-543-681A-6465 Sequence 6465, Ap
43 97 5.7 541 4 US-09-134-000C-5420 Sequence 5420, Ap
44 97 5.7 2482 1 US-08-328-254-6 Sequence 6, Appli
45 97 5.7 3248 1 US-08-353-700-1 Sequence 1, Appli

```

#### ALIGNMENTS

RESULT 1

US-09-190-965-1

; Sequence 1, Application US/09190965

; Patent No. 6071721

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/09/190,965

; CURRENT FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 1

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 3734805

US-09-190-965-1

```

Query Match      100.0%; Score 1704; DB 3; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.7e-161;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MKKMPFLFSKHNPAIVKILKDNLAILEKQDKTDKASEEVSLSQAMKEILCGTNEKE 60
DB 1 MKKMPFLFSKHNPAIVKILKDNLAILEKQDKTDKASEEVSLSQAMKEILCGTNEKE 60
QY 61 PTEAQAQAEQLYSGLLVTLIADQLDLDFEGKDVTOIFNNILRQIGTRSPTEYIS 120
DB 61 PTEAQAQAEQLYSGLLVTLIADQLDLDFEGKDVTOIFNNILRQIGTRSPTEYIS 120
QY 121 AHPHILFMLKGYEAPQIALRCGIMRECIRHEPLAKIILFSNQFRDFFKYVELSTFDIA 180
DB 121 AHPHILFMLKGYEAPQIALRCGIMRECIRHEPLAKIILFSNQFRDFFKYVELSTFDIA 180
QY 181 SDAFATFKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENVTYTKRQSLKLGELIDRH 240
DB 181 SDAFATFKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENVTYTKRQSLKLGELIDRH 240
QY 241 NFAIMTKYISKPENLKMNNLRDPSNIOFAHFVKVFVSPHKTQPIVEILLKNQPK 300
DB 241 NFAIMTKYISKPENLKMNNLRDPSNIOFAHFVKVFVSPHKTQPIVEILLKNQPK 300
QY 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337
DB 301 LIEFLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337

```

```
RESULT 2
US-09-470-253-1
; Sequence 1, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR FILING DATE: 1999-12-22
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-470-253-1

Query Match      100.0%; Score 1704; DB 4; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.7e-161;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKPLFSKSHKNPAEIVKILKDNLAILEKQDIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120
DB 1 MKKPLFSKSHKNPAEIVKILKDNLAILEKQDIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120

QY 61 PPTAEVAQLAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120
DB 61 PPTAEVAQLAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120

QY 121 APHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDIA 180
DB 121 APHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDIA 180

QY 181 SDAPATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDRH 240
DB 181 SDAPATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDRH 240

QY 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 300
DB 241 NFAIMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 300

QY 301 LIBFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 337
DB 301 LIBFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 337

RESULT 3
US-09-190-965-3
; Sequence 3, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-470-253-3

Query Match      80.8%; Score 1376; DB 4; Length 341;
Best Local Similarity 80.7%; Pred. No. 8e-129;
Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;

QY 4 MPL-FSKSHKNPAEIVKILKDNLAILEKQDIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 119
DB 1 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKKAERKATEEVSKNLVAMKEILYGTNEK 60

QY 60 EPTAEVAQLAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 119
DB 61 EPQTEAVALAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120

QY 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDI 179
DB 121 CTQQNILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDI 180

QY 180 ASDAFATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDR 239
DB 181 ASDAFATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDR 240

QY 240 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 299
DB 241 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 300

QY 300 KLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 336
DB 301 KLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 337

RESULT 4
US-09-470-253-3
; Sequence 3, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-470-253-3

Query Match      80.8%; Score 1376; DB 4; Length 341;
Best Local Similarity 80.7%; Pred. No. 8e-129;
Matches 272; Conservative 32; Mismatches 29; Indels 4; Gaps 2;

QY 4 MPL-FSKSHKNPAEIVKILKDNLAILEKQDIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 119
DB 1 MPFPFGKSHKSPADIVKLNKESMAVLEKQDISDKKAERKATEEVSKNLVAMKEILYGTNEK 60

QY 60 EPTAEVAQLAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 119
DB 61 EPQTEAVALAQLYSSGLLVTLADIQLIDFEKGDVTOIFNNILRRQIGTRSPVVEYS 120

QY 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDI 179
DB 121 CTQQNILFMLKGYEAPQIALRCGIMLRCEIRHPEPLAKIILFSNQFRDFFKYVELSTFDI 180

QY 180 ASDAFATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDR 239
DB 181 ASDAFATFKDLLTRHKLVADFLQNYDTIFEDYEKLQSENYYVTKRQSLKLGELILDR 240

QY 240 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 299
DB 241 HNFATMTKYISKPENLKMNNLLRDKSPNIOFEAFHVKVFPVSPHKTQPIVEILLKNQPK 300

QY 300 KLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 336
DB 301 KLIEFLSSFOKERTDDEQFADEKKNYLIKQIRDLKKTAP 337
```

Db 121 CTQONILFMLLKGYSFELALNGIMLRECIHREPLAKILMSEQYDFRIVEMSTFDI 180  
QY 180 ASDAFATPKDLLTRHKVLVADLEQNYDTIFEDYEKLLQSENVTYKRSQSLKLGELILDR 239  
Db 181 ASDAFATPKDLLTRHKVLSAEFLQHYDRFFSEYKLLHSENVTYKRSQSLKLGELILDR 240  
QY 240 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 299  
Db 241 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 300  
QY 300 KLIEFLSSFOKERTDDQFADEKKNYLIKQIRDLKTA 336  
Db 301 KLIEFLSKQNDTEDEQFNDEKTYLVKQIRNLKRAA 337

## RESULT 5

US-09-190-965-4  
; Sequence 4, Application US/09190965  
; Patent No. 6071721

; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Neil J.  
; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program  
; SEQ ID NO 4

; LENGTH: 339  
; TYPE: PRT

; ORGANISM: Drosophila melanogaster  
; FEATURE: -

; OTHER INFORMATION: g1794137  
US-09-190-965-4

Query Match 65.1%; Score 1109; DB 3; Length 339;

Best Local Similarity 65.0%; Pred. No. 3.1e-102;

Matches 217; Conservative 59; Mismatches 54; Indels 4; Gaps 3;

QY 4 MPEFSKSHKNPAEIVKILKDNLAILEKQDKTKDASEVSKSLOAMKEILCGTNEKEPPT 63  
Db 1 MPEFGKSQSPVELVKSLEKAINALEAGDRKVEKAQEDVSKNLVSKNMLHSSDAEPPA 60

QY 64 E-AVAQLAQELYSGLLVTLIADLQIDFEGKKDVTFQFNILRRQIGTRSPTEVEISAH 122  
Db 61 DYVVAQLSQELYNLLLLLQNLHRIDFEGKKHVALIFNNLLRRQIGTRSPTEVEICTK 120

QY 123 PHILFMLLKGYE--APQIALRGIMLRECIHREPLAKILIFSNQFDFKYYVELSTFDIA 180  
Db 121 PEILFTMAGYEDAHPEIALNSGTMLRECARYEALAKIMLHSDFFKFFRYVEVSTFDIA 180

QY 181 SDAFATPKDLLTRHKVLVADFLSQNYDTIF-EDYEKLLQSENVTYKRSQSLKLGELILDR 239  
Db 181 SDAFSTFKELLTRHKVLCAEFLDANYDFEFSQHYQELNSENVTYKRSQSLKLGELILDR 240

QY 240 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 299  
Db 241 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 300

QY 300 KLIEFLSSFOKERTDDQFADEKKNYLIKQIRDLK 333  
Db 301 KLVDFLTNFHTDRSEDEQFNDEKAYLIKQIKELK 334

## RESULT 6

US-09-470-253-4

; Sequence 4, Application US/09470253  
; Patent No. 6365371

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Neil J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470,253  
; PRIOR FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/190,965  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 339  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE: -  
; OTHER INFORMATION: g1794137  
US-09-470-253-4

Query Match 65.1%; Score 1109; DB 4; Length 339;

Best Local Similarity 65.0%; Pred. No. 3.1e-102;

Matches 217; Conservative 59; Mismatches 54; Indels 4; Gaps 3;

QY 4 MPEFSKSHKNPAEIVKILKDNLAILEKQDKTKDASEVSKSLOAMKEILCGTNEKEPPT 63  
Db 1 MPEFGKSQSPVELVKSLEKAINALEAGDRKVEKAQEDVSKNLVSKNMLHSSDAEPPA 60

QY 64 E-AVAQLAQELYSGLLVTLIADLQIDFEGKKDVTFQFNILRRQIGTRSPTEVEISAH 122  
Db 61 DYVVAQLSQELYNLLLLLQNLHRIDFEGKKHVALIFNNLLRRQIGTRSPTEVEICTK 120

QY 123 PHILFMLLKGYE--APQIALRGIMLRECIHREPLAKILIFSNQFDFKYYVELSTFDIA 180  
Db 121 PEILFTMAGYEDAHPEIALNSGTMLRECARYEALAKIMLHSDFFKFFRYVEVSTFDIA 180

QY 181 SDAFATPKDLLTRHKVLVADFLSQNYDTIF-EDYEKLLQSENVTYKRSQSLKLGELILDR 239  
Db 181 SDAFSTFKELLTRHKVLCAEFLDANYDFEFSQHYQELNSENVTYKRSQSLKLGELILDR 240

QY 240 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 299  
Db 241 HNFATMTKYISKPENLKLMMNLLRDKSPNIOFEAFHVKFVFNASPHKTOPIVEILLKNOP 300

QY 300 KLIEFLSSFOKERTDDQFADEKKNYLIKQIRDLK 333  
Db 301 KLVDFLTNFHTDRSEDEQFNDEKAYLIKQIKELK 334

## RESULT 7

US-09-190-965-5

; Sequence 5, Application US/09190965

; Patent No. 6071721

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Neil J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 5

; LENGTH: 377

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

; FEATURE: -

; OTHER INFORMATION: g1255838

; US-09-190-965-5

Query Match

62.4%; Score 1063.5; DB 3; Length 377;

Best Local Similarity 60.5%; Pred. No. 1.2e-97;  
Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;  
QY 4 MP-LFSKSHKNPAEIVKILKDNLAILEK-----QDKTDXKASEEVSLSQAM 49  
Db 1 MPLLFGKSHKSPADVKTLLREVLITLTKLPPPKLDKDGNIQSDKKYDKALDEVSKNVAMI 60  
QY 50 KEILCGTNEKEPTE---AVAQAEYLYSSGLVTLIADLQIDFPGKDVTOIFNNILR 106  
Db 61 KSFYIGNDSAPSESEHVQVAQAEYVYNANILPMLIKMLPKFEPCCKDVGVQIFNNLLR 120  
QY 107 RQIGTRSPTEYISAHPHILFMLLKGYEAPQIALRCGIMLRCIRHEPLAKIILFNSQFR 166  
Db 121 RQIGTRSPTEYILGAPPEILLQVQYGVSDPIALTGMLRESIRHDLAKIILYSDVY 180  
QY 167 DFFKYVELSTFDIASDAFATFKDOLLTRHKVVLVADFLQNYDTIFEDYEKLQSENVTYTR 226  
Db 181 TFFLYVQSEVFDISSDAFSTFKELTTRHKAIIEFLDSNYDTFFAQYQNLNSKNVYTR 240  
QY 227 QSLKLGELILDRHNPATMYKISKPENLKMNNLLRDKSPNTQPEAFHVKFVVASPHK 286  
Db 241 QSLKLGELILDRHNPATMYKISNPDLNLMELLRDKSRNTQYEAFFHVKFVVASPNK 300  
QY 287 TQPIVEILLKNQPKLIEFLSSQKERTDDQFADEKNYLIKQIRDLKKT 335  
Db 301 PKPISDILNREKLVSEFSEHNDRTDDEQFNDEKAYLIKQIQEMKSS 349

## RESULT 8

US-09-470-253-5  
; Sequence 5, Application US/09470253  
; Patent No. 6365371  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470, 253  
; PRIOR FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/130,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
; FEATURE: -  
; OTHER INFORMATION: g1255838  
US-09-470-253-5

Query Match 62.4%; Score 1063.5; DB 4; Length 377;  
Best Local Similarity 60.5%; Pred. No. 1.2e-97;  
Matches 211; Conservative 53; Mismatches 68; Indels 17; Gaps 3;  
QY 4 MP-LFSKSHKNPAEIVKILKDNLAILEK-----QDKTDXKASEEVSLSQAM 49  
Db 1 MPLLFGKSHKSPADVKTLLREVLITLTKLPPPKLDKDGNIQSDKKYDKALDEVSKNVAMI 60  
QY 50 KEILCGTNEKEPTE---AVAQAEYLYSSGLVTLIADLQIDFPGKDVTOIFNNILR 106  
Db 61 KSFYIGNDSAPSESEHVQVAQAEYVYNANILPMLIKMLPKFEPCCKDVGVQIFNNLLR 120  
QY 107 RQIGTRSPTEYISAHPHILFMLLKGYEAPQIALRCGIMLRCIRHEPLAKIILFNSQFR 166  
Db 121 RQIGTRSPTEYILGAPPEILLQVQYGVSDPIALTGMLRESIRHDLAKIILYSDVY 180  
QY 167 DFFKYVELSTFDIASDAFATFKDOLLTRHKVVLVADFLQNYDTIFEDYEKLQSENVTYTR 226  
Db 181 TFFLYVQSEVFDISSDAFSTFKELTTRHKAIIEFLDSNYDTFFAQYQNLNSKNVYTR 240

QY 227 QSLKLGELILDRHNPATMYKISKPENLKMNNLLRDKSPNTQPEAFHVKFVVASPHK 286  
Db 241 QSLKLGELILDRHNPATMYKISNPDLNLMELLRDKSRNTQYEAFFHVKFVVASPNK 300  
QY 287 TQPIVEILLKNQPKLIEFLSSQKERTDDQFADEKNYLIKQIRDLKKT 335  
Db 301 PKPISDILNREKLVSEFSEHNDRTDDEQFNDEKAYLIKQIQEMKSS 349  
RESULT 9  
US-09-914-259-11  
; Sequence 11, Application US/09914259  
; Patent No. 6495336  
; GENERAL INFORMATION:  
; APPLICANT: Makowski, Lee  
; APPLICANT: Hyman, Paul  
; APPLICANT: Williams, Mark  
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES  
; FILE REFERENCE: 8471-010-999  
; CURRENT APPLICATION NUMBER: US/09/914, 259  
; CURRENT FILING DATE: 2000-11-21  
; NUMBER OF SEQ ID NOS: 180  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 3878  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-914-259-11

Query Match 7.5%; Score 128.5; DB 4; Length 3878;  
Best Local Similarity 20.1%; Pred. No. 0.0044;  
Matches 77; Conservative 75; Mismatches 116; Indels 115; Gaps 15;  
QY 18 VKILKDNLAILEKQDKTDXKASEEVSLSQAMKEILCGTNEKEPTEAVAQAEYLYSSG 77  
Db 664 IEKLDKDLGIHVQK--QIDGLQNMESQKTIETMQ-----FEKDNLTITKQNLILE----- 710  
QY 78 LLVTLIADLQ--LIDPEGKDVTOIFNNILRQI-----GTRSPTEYISAHPHI 125  
Db 711 --ISKLDLQQLSVNSKSEMTLQI--NELQKEILROEKEKEKGLTQEVQELQKTEL 766  
QY 126 LFMLLKGYEAPQIALRCGIMLRCIRHEPLAKIILFNSQFRDFFKYVELSTFDIASDAFA 185  
Db 767 LEQMEKE-----NDLQEKFAQLEAEN--SILKDEK 797  
QY 186 TPKDILLTRH-----KVLVADFLQ--QNYDTIFEDYEKLQSENVTYTRKQSLKLGELIL 237  
Db 798 TLEDMLKIHTPVQSEERLIFLDSIKSKSDSVWEKEIILIBENEDLKQCCIQALNEIEK 857  
QY 238 DRHNPATMYK-----YISKPENLKMNNLLRD 264  
Db 858 QRNTFSFAEKNFEVNTQOEYACLLKVKDDLSKKNQOELEYKSKLKAINEELHLQRI 917  
QY 265 KSPNIQFEA--FHVFKVVASPHKTPQIVEILLKNQPKLIEFLSSQKERTD--DEQFAD-- 320  
Db 918 NPTTVKMKSSVFDEDTFVA---ETLEMGVVEKDTTLMEXLEVTREKLELSQRLSDL 974  
QY 321 -----EKVLIKQIRDLK 334  
Db 975 SEQLKQKHGEISPLNEEVSKLQ 997  
RESULT 10  
US-09-724-517-2  
; Sequence 2, Application US/09724517  
; Patent No. 6379941  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6379941el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1031  
; CURRENT APPLICATION NUMBER: US/09/724, 517

; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US/09/641,807  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1279  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (409)...(436)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-724-517-2

Query Match 6.7%; Score 113.5; DB 4; Length 1279;  
Best Local Similarity 19.3%; Pred. No. 0.028;  
Matches 87; Conservative 61; Mismatches 137; Indels 165; Gaps 14;  
QY 23 DNLAILEKQDKKTDKASEEVSLSQAMKEILCGTNEKEPPTTEAVQAQELYSGLLVT 82  
DB 794 DHLQKLDQKKWLDVEVEKVLNQRQLELEADLKREAIIVSKKEALLQE--KSHLENKK 851  
QY 83 IADLQIDFEGKKDVTQIFNNILRRQIGTRSPTEYISA----- 121  
DB 852 LRSSQALNTDSLKISTR--NLLEQELSEKNVLOTSTAEKTKISEQVEVLQKEKDQLQ 909  
QY 122 -----HPHILFMLKGYEAPQIALRCGIMLRECIHRHEPLAKIILFS 162  
DB 910 KRRHDVDEKLNKGRVLSPEEHLVFLQEEGIEALEAAIE---YRNESIQNRQKSLRASFS 966  
QY 163 NQFRDFFKVE---LSTFDIASDAFATFKDILT-----RHKVLVAD--- 200  
DB 967 NLSRGEANVLEKACLSPVEITILFRYFNKVNLRREARKQOQLYNEEMKVKLDRDNMV 1026  
QY 201 -----FLEQNYDTI-----FEDYEKLLQS 219  
DB 1027 RELESALDHLKQCDRLTLQOKEHEQKQMLLHHFKEDQGGIMETFTYEDKIQOLEK 1086  
QY 220 ENYVTKROS-----LKLGLILDRHNFAM-----TKYISK 251  
DB 1087 DLYFYKTSRDHKKKLKELVGEAI--RRQLAPSEYQEAGDGVLPKGGMLSEELKWA 1144  
QY 252 PENLKLMMNLLRDKSPNIQFEAFHVKFVSPHKTQPIVEILLKNQPKLIEFLSSF--- 308  
DB 1145 PESMKLSG---REREMDSS-----ASLRTQPNQPKLWEDIPELPIHSSLAPP 1190  
QY 309 -----QKERTDDEQFADEKNYLIKQIR 330  
DB 1191 SGHMLGNENKTETDDNQFTKSHRSLSSQIQ 1220

RESULT 11  
US-09-641-807A-2  
; Sequence 2, Application US/09641807A  
; Patent No. 6440731  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6440731el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1031  
; CURRENT APPLICATION NUMBER: US/09/641,807A  
; CURRENT FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1279  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (409)...(446)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-724-517-2

; OTHER INFORMATION: Xaa = any amino acid  
US-09-641-807A-2  
Query Match 6.7%; Score 113.5; DB 4; Length 1279;  
Best Local Similarity 19.3%; Pred. No. 0.028;  
Matches 87; Conservative 61; Mismatches 137; Indels 165; Gaps 14;  
QY 23 DNLAILEKQDKKTDKASEEVSLSQAMKEILCGTNEKEPPTTEAVQAQELYSGLLVT 82  
DB 794 DHLQKLDQKKWLDVEVEKVLNQRQLELEADLKREAIIVSKKEALLQE--KSHLENKK 851  
QY 83 IADLQIDFEGKKDVTQIFNNILRRQIGTRSPTEYISA----- 121  
DB 852 LRSSQALNTDSLKISTR--NLLEQELSEKNVLOTSTAEKTKISEQVEVLQKEKDQLQ 909  
QY 122 -----HPHILFMLKGYEAPQIALRCGIMLRECIHRHEPLAKIILFS 162  
DB 910 KRRHDVDEKLNKGRVLSPEEHLVFLQEEGIEALEAAIE---YRNESIQNRQKSLRASFS 966  
QY 163 NQFRDFFKVE---LSTFDIASDAFATFKDILT-----RHKVLVAD--- 200  
DB 967 NLSRGEANVLEKACLSPVEITILFRYFNKVNLRREARKQOQLYNEEMKVKLDRDNMV 1026  
QY 201 -----FLEQNYDTI-----FEDYEKLLQS 219  
DB 1027 RELESALDHLKQCDRLTLQOKEHEQKQMLLHHFKEDQGGIMETFTYEDKIQOLEK 1086  
QY 220 ENYVTKROS-----LKLGLILDRHNFAM-----TKYISK 251  
DB 1087 DLYFYKTSRDHKKKLKELVGEAI--RRQLAPSEYQEAGDGVLPKGGMLSEELKWA 1144  
QY 252 PENLKLMMNLLRDKSPNIQFEAFHVKFVSPHKTQPIVEILLKNQPKLIEFLSSF--- 308  
DB 1145 PESMKLSG---REREMDSS-----ASLRTQPNQPKLWEDIPELPIHSSLAPP 1190  
QY 309 -----QKERTDDEQFADEKNYLIKQIR 330  
DB 1191 SGHMLGNENKTETDDNQFTKSHRSLSSQIQ 1220

RESULT 12  
US-09-723-096-2  
; Sequence 2, Application US/09723096  
; Patent No. 6448026  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6448026el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1031  
; CURRENT APPLICATION NUMBER: US/09/723,096  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US/09/641,807  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1279  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (409)...(436)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-723-096-2

Query Match 6.7%; Score 113.5; DB 4; Length 1279;  
Best Local Similarity 19.3%; Pred. No. 0.028;  
Matches 87; Conservative 61; Mismatches 137; Indels 165; Gaps 14;  
QY 23 DNLAILEKQDKKTDKASEEVSLSQAMKEILCGTNEKEPPTTEAVQAQELYSGLLVT 82  
DB 794 DHLQKLDQKKWLDVEVEKVLNQRQLELEADLKREAIIVSKKEALLQE--KSHLENKK 851

RESULT 15  
US-08-630-822A-70  
; Sequence 70, Application US/08630822A  
; Patent No. 5840695  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.  
; APPLICANT: HUNTER, SHIRLEY WU  
; APPLICANT: WALLENFELS, LYNDIA  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS

Search completed: April 12, 2004, 10:32:58  
Job time : 24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:33:03 ; Search time 20 Seconds

(without alignments)  
1620.827 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 337

Sequence: 1 MKKMPLFSKSHKPAEIVKI.....FADEKNYLKQIRDLKKTAP 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

PIR 78: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	6.8	341	2	157997
2	16	4.7	329	2	TS0117
3	14	4.2	338	2	T27129
4	14	4.2	377	2	T16651
5	8	2.4	923	2	S03900
6	7	2.1	103	2	S72925
7	7	2.1	122	2	G89814
8	7	2.1	134	2	AG2552
9	7	2.1	135	2	AH2522
10	7	2.1	135	2	F71407
11	7	2.1	154	2	AB1605
12	7	2.1	156	2	C64300
13	7	2.1	171	2	G95316
14	7	2.1	179	2	G70325
15	7	2.1	191	2	AB2247
16	7	2.1	229	2	SI9734
17	7	2.1	241	2	T27075
18	7	2.1	264	2	T23866
19	7	2.1	275	2	C87153
20	7	2.1	275	2	D70896
21	7	2.1	282	2	E96937
22	7	2.1	283	2	G98020
23	7	2.1	299	2	H89906
24	7	2.1	301	2	T40402
25	7	2.1	305	2	G71441
26	7	2.1	309	1	E64762
27	7	2.1	309	2	D90679
28	7	2.1	309	2	H85529
29	7	2.1	309	2	T17557
30	7	2.1	314	2	T43132
31	7	2.1	320	2	F70034
32	7	2.1	332	2	G90291
33	7	2.1	334	2	G90360
34	7	2.1	334	2	T08491
35	7	2.1	334	2	T08494
36	7	2.1	336	2	S04682
37	7	2.1	340	1	PEIKL
38	7	2.1	344	2	B96987
39	7	2.1	357	2	G72746
40	7	2.1	359	2	AD3559
41	7	2.1	365	2	F82210
42	7	2.1	368	2	G96668
43	7	2.1	372	2	I64223
44	7	2.1	373	2	G70355
45	7	2.1	376	2	T16059
46	7	2.1	381	2	T33761
47	7	2.1	399	2	S34681
48	7	2.1	400	2	G95936
49	7	2.1	405	2	S73853
50	7	2.1	421	2	D95975
51	7	2.1	430	2	D96719
52	7	2.1	436	2	F97058
53	7	2.1	440	2	E90563
54	7	2.1	444	2	JC4348
55	7	2.1	474	2	S46175
56	7	2.1	490	2	JE0096
57	7	2.1	504	2	JC5830
58	7	2.1	521	2	D82508
59	7	2.1	523	2	D85538
60	7	2.1	523	2	H90687
61	7	2.1	543	2	T27190
62	7	2.1	552	2	T27191
63	7	2.1	554	2	F86244
64	7	2.1	576	2	A84902
65	7	2.1	587	2	T24103
66	7	2.1	624	2	T44840
67	7	2.1	731	2	B86369
68	7	2.1	758	2	B82122
69	7	2.1	762	2	T38081
70	7	2.1	786	2	AG2375
71	7	2.1	818	2	AG7668
72	7	2.1	818	2	AF2892
73	7	2.1	831	2	A70363
74	7	2.1	846	2	AD0279
75	7	2.1	923	1	MMBY7C
76	7	2.1	962	2	C43274
77	7	2.1	1055	2	A87364
78	7	2.1	1115	2	A45761
79	7	2.1	1170	2	A72287
80	7	2.1	1239	2	I49705
81	7	2.1	1250	1	B45219
82	7	2.1	1292	2	D4727
83	7	2.1	1489	2	T38842
84	7	2.1	2013	2	C71610
85	7	2.1	2829	2	A42771
86	7	2.1	3433	1	S28381
87	7	2.1	3788	2	T30851
88	7	2.1	3788	2	T31960
89	6	1.8	15	2	JP0101
90	6	1.8	25	2	B41606
91	6	1.8	40	2	G45495
92	6	1.8	40	2	I45495
93	6	1.8	46	2	PC4162
94	6	1.8	61	2	A46257
95	6	1.8	61	2	T37147
96	6	1.8	67	2	E95089
97	6	1.8	72	2	G71355
98	6	1.8	73	2	H97956
99	6	1.8	75	2	T42944
100	6	1.8	75	2	A64600
101	6	1.8	80	2	AF2836
102	6	1.8	82	2	T10329

hypothetical prote  
conserved hypothet  
endoglucanase prec  
endoglucanase prec  
probable transposa  
probable transposa  
ribosomal protein  
polyporopepsin (EC  
probable transcrip  
probable carbamoyl  
ornithine cyclodea  
amino acid ABC tra  
protein FIN19.7 [i  
bifunctional endo-  
8-amino-7-oxononan  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical outer  
Fe-S oxidoreductas  
virulence-mediati  
probable resistanc  
myocillin - mouse  
myocillin - human  
methyl-accepting c  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
auxin-regulated pr  
hypothetical prote  
probable dTPGluco  
hypothetical prote  
hypothetical prote  
ferrous iron trans  
probable mitochon  
WD-40 repeat-pro  
mannosidase AGR C  
mannosidase (impor  
mannose-1-phosphat  
probable virulence  
probable membrane  
N-methyl D-asparta  
OmpA-related prote  
Ca2+-transporting  
hypothetical prote  
glutamate receptor  
N-methyl-D-asparta  
probable RAD50 DNA  
probable RAS GTPas  
probable membrane  
reticulocyte-bind  
utrophin - human  
lysosomal traffic  
beige protein homo  
fibrinogen alpha c  
homeotic protein M  
beta-defensin-7 -  
beta-defensin-9 -  
toxin-co-regulated  
Dbx homeobox (home  
hypothetical prote  
hypothetical prote  
probable ribosomal  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote

103	6	1.8	83	2	S75083	hypothetical prote	176	6	1.8	150	2	B36848	All protein - vari
104	6	1.8	87	2	A64306	ribosomal protein	177	6	1.8	150	2	T28542	hypothetical prote
105	6	1.8	88	2	C86200	hypothetical prote	178	6	1.8	150	2	T37387	VLRF-2, late gene
106	6	1.8	88	2	E71896	hypothetical prote	179	6	1.8	150	2	T22617	hypothetical prote
107	6	1.8	89	2	G90777	hypothetical prote	180	6	1.8	150	2	T23641	hypothetical prote
108	6	1.8	90	2	A27056	erythrocyte membra	181	6	1.8	150	2	D37188	uncharacterized pr
109	6	1.8	90	2	D64617	hypothetical prote	182	6	1.8	150	2	A55209	H transfer determi
110	6	1.8	93	2	AH1987	hypothetical prote	183	6	1.8	150	2	AC1758	protein gp15 from
111	6	1.8	95	2	F95332	hypothetical prote	184	6	1.8	151	2	AE2193	hypothetical prote
112	6	1.8	97	2	AE2631	hypothetical prote	185	6	1.8	152	2	G71820	hypothetical prote
113	6	1.8	97	2	G97413	hypothetical prote	186	6	1.8	152	2	D64700	small protein - He
114	6	1.8	98	2	AI2262	hypothetical prote	187	6	1.8	152	2	AI3529	flbr protein limpo
115	6	1.8	99	2	F82351	gene 3 protein-rel	188	6	1.8	152	2	AI3529	conserved hypothet
116	6	1.8	100	2	H70413	NADH2 dehydrogenas	189	6	1.8	153	2	AI0438	replication-associ
117	6	1.8	101	1	B44056	ilk protein - cani	190	6	1.8	155	2	T10115	interleukin-2 prec
118	6	1.8	103	2	A38354	carboxypeptidase B	191	6	1.8	155	2	S11488	interleukin-2 prec
119	6	1.8	107	4	A60600	hypothetical mutan	192	6	1.8	156	2	AC0467	conserved hypothet
120	6	1.8	108	2	E72494	hypothetical prote	193	6	1.8	157	2	T24304	hypothetical prote
121	6	1.8	108	2	AF0091	hypothetical prote	194	6	1.8	157	2	F87509	conserved hypothet
122	6	1.8	109	2	A83265	conserved hypothet	195	6	1.8	157	2	D90045	hypothetical prote
123	6	1.8	109	2	A89990	hypothetical prote	196	6	1.8	157	2	AG1312	thioredoxin homolo
124	6	1.8	110	2	A24444	hypothetical prote	197	6	1.8	157	2	AG1684	thioredoxin homolo
125	6	1.8	110	2	T12499	hypothetical prote	198	6	1.8	158	2	D70438	hypothetical prote
126	6	1.8	112	2	S74015	hypothetical prote	199	6	1.8	159	2	C65113	yhvE protein - Esc
127	6	1.8	113	2	S23653	sensorin A - Calif	200	6	1.8	160	2	D72226	conserved hypothet
128	6	1.8	122	2	C86336	hypothetical prote	201	6	1.8	161	2	S73630	ribosomal protein
129	6	1.8	123	2	A99504	hypothetical prote	202	6	1.8	161	2	AI2638	phas protein limpo
130	6	1.8	124	1	A69363	conserved hypothet	203	6	1.8	162	2	I64239	ribosomal protein
131	6	1.8	124	2	F83560	probable type II s	204	6	1.8	162	2	B97421	probable poly(3-hy
132	6	1.8	128	2	T06966	hypothetical prote	205	6	1.8	164	2	E75100	hypothetical prote
133	6	1.8	128	2	A81153	type I restriction	206	6	1.8	165	2	G72486	hypothetical prote
134	6	1.8	129	2	H71046	hypothetical prote	207	6	1.8	167	2	AB2596	biopolymer transpo
135	6	1.8	129	2	H90161	conserved hypothet	208	6	1.8	167	2	F97378	biopolymer transpo
136	6	1.8	129	2	S48814	hypothetical prote	209	6	1.8	168	2	S42070	early chorion prot
137	6	1.8	130	2	H72722	hypothetical prote	210	6	1.8	168	2	B96916	hypothetical prote
138	6	1.8	131	2	D84401	30S ribosomal prot	211	6	1.8	169	2	S23060	early chorion prot
139	6	1.8	131	2	T25924	sugar transport ho	212	6	1.8	169	2	S42071	early chorion prot
140	6	1.8	132	2	T30595	ribosomal protein	213	6	1.8	169	2	C90175	NADH dehydrogenase
141	6	1.8	133	2	D72054	S8 ribosomal prote	214	6	1.8	170	2	F55548	transcription regu
142	6	1.8	133	2	G86569	hypothetical prote	215	6	1.8	171	2	D81278	probable periplasm
143	6	1.8	133	2	F70334	hypothetical prote	216	6	1.8	172	2	S27022	fibroblast growth
144	6	1.8	134	2	S26612	ribosomal protein	217	6	1.8	172	2	G64555	hypothetical prote
145	6	1.8	134	2	F72638	hypothetical prote	218	6	1.8	173	2	F81122	hypothetical prote
146	6	1.8	135	2	B69103	translation initia	219	6	1.8	173	2	F70112	hypothetical prote
147	6	1.8	135	2	E69842	hypothetical prote	220	6	1.8	174	2	A35383	superoxide dismuta
148	6	1.8	136	2	B90443	hypothetical prote	221	6	1.8	174	2	F75606	hypothetical prote
149	6	1.8	137	2	T17944	hypothetical prote	222	6	1.8	176	2	E90006	hypothetical prote
150	6	1.8	137	2	T07027	hypothetical prote	223	6	1.8	176	2	T02217	NBS-LRR type resis
151	6	1.8	140	2	E75056	probable translati	224	6	1.8	178	2	AB0414	hypoxanthine phosp
152	6	1.8	140	2	D71104	probable translati	225	6	1.8	178	2	AB0523	hypoxanthine phosp
153	6	1.8	140	2	S34667	hypothetical prote	226	6	1.8	178	2	T23601	hypothetical prote
154	6	1.8	141	1	HASH	hemoglobin alpha c	227	6	1.8	180	2	E70358	HuPS hydrogense r
155	6	1.8	142	2	I48552	orf US426 - infect	228	6	1.8	181	2	AI0861	syd protein (impor
156	6	1.8	143	1	HKW41	heat shock protein	229	6	1.8	182	2	T07837	thioredoxin f prec
157	6	1.8	144	2	T15047	RNA binding protei	230	6	1.8	182	2	S20929	thioredoxin f prec
158	6	1.8	144	2	T16961	probable membrane	231	6	1.8	182	2	A90645	hypoxanthine phosp
159	6	1.8	144	2	AI0008	ribosomal protein	232	6	1.8	182	2	A85496	hypoxanthine phosp
160	6	1.8	145	2	B82878	hypothetical prote	233	6	1.8	182	2	S45202	hypoxanthine phosp
161	6	1.8	145	2	G69011	hypothetical prote	234	6	1.8	182	2	F71072	hypothetical prote
162	6	1.8	145	2	G75031	hypothetical prote	235	6	1.8	182	2	F70509	hypothetical prote
163	6	1.8	145	2	E29010	mer operon ORF2 hy	236	6	1.8	183	2	G71921	hypothetical prote
164	6	1.8	146	2	T10511	hypothetical prote	237	6	1.8	183	2	C85439	probable cytoskele
165	6	1.8	147	2	E90104	40S ribosomal prot	238	6	1.8	184	2	C84592	hypothetical prote
166	6	1.8	147	2	AD1736	conserved hypothet	239	6	1.8	185	2	S27936	hypothetical prote
167	6	1.8	148	2	B71137	hypothetical prote	240	6	1.8	186	1	JQ1623	envelope-associate
168	6	1.8	148	2	B85641	hypothetical prote	241	6	1.8	186	2	JQ1987	hypothetical 21.0K
169	6	1.8	148	2	C90780	hypothetical prote	242	6	1.8	188	2	F83816	late competence op
170	6	1.8	148	2	AE0928	probable phage tai	243	6	1.8	189	2	AG3395	NADH2 dehydrogenas
171	6	1.8	148	2	D90143	conserved hypothet	244	6	1.8	190	2	G89581	protein Cl8Al1.4 [
172	6	1.8	149	2	AI2048	hypothetical prote	245	6	1.8	191	2	G70940	hypothetical prote
173	6	1.8	150	1	WMV2R2	17K protein - vacc	246	6	1.8	191	2	C81279	probable molybdopt
174	6	1.8	150	2	F72163	All protein - vari	247	6	1.8	192	2	S76506	hypothetical prote
175	6	1.8	150	2	C42517	All protein - vacc	248	6	1.8	192	2	AD2488	hypothetical prote

249 1.8 195 2 A84191  
250 1.8 199 2 JC7106  
251 1.8 200 2 H82570  
252 1.8 200 2 C81282  
253 1.8 201 2 S57135  
254 1.8 201 2 C85078  
255 1.8 201 2 S52833  
256 1.8 202 2 D83578  
257 1.8 202 2 T49823  
258 1.8 203 2 S18684  
259 1.8 203 2 S43222  
260 1.8 203 2 H90454  
261 1.8 206 2 AC2301  
262 1.8 206 2 T16153  
263 1.8 206 2 S63540  
264 1.8 207 1 S06869  
265 1.8 207 2 E89917  
266 1.8 208 1 B69066  
267 1.8 209 2 F90348  
268 1.8 210 2 S77362  
269 1.8 210 2 S76008  
270 1.8 210 2 G95228  
271 1.8 210 2 T41982  
272 1.8 210 2 C98093  
273 1.8 211 1 ZK9PT4  
274 1.8 212 1 G64331  
275 1.8 212 2 F64122  
276 1.8 213 2 G95079  
277 1.8 213 2 B97947  
278 1.8 214 2 JC7297  
279 1.8 215 2 H71451  
280 1.8 215 2 D75215  
281 1.8 217 2 H83351  
282 1.8 217 2 B64445  
283 1.8 217 2 T04580  
284 1.8 217 2 B97119  
285 1.8 218 2 H64091  
286 1.8 218 2 AG1725  
287 1.8 219 2 D81792  
288 1.8 220 2 E82364  
289 1.8 220 2 AG2035  
290 1.8 222 1 CEECFE  
291 1.8 222 2 AD0464  
292 1.8 222 2 A00991  
293 1.8 222 2 S25014  
294 1.8 222 2 H91167  
295 1.8 222 2 H86013  
296 1.8 223 2 A55847  
297 1.8 223 2 F83598  
298 1.8 223 2 E96955  
299 1.8 224 2 F71678  
300 1.8 224 2 S48671

## ALIGNMENTS

RESULT 1  
157997  
Hypothetical calcium-binding protein - mouse  
C:Species: Mus sp. (mouse)  
C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 19-May-2000  
C:Accession: 157997  
R.Wiyamoto, H.; Matsuhiro, A.; Nozaki, M.  
Mol. Reprod. Dev. 34, 1-7, 1993  
A>Title: Molecular cloning of a novel mRNA sequence expressed in cleavage stage mouse embryo  
A:Reference number: 157997, MUID:93119656; PMID:8418809  
A:Accession: 157997  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-341 <RES>  
A:Cross-references: GB:S51858; NID:G262933; PIDN:AAB24801.1; PID:G262934  
C:Superfamily: Saccharomyces hypothetical protein YKL189w

C:Keywords: calcium binding  
Query Match 6.8%; Score 23; DB 2; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1e-15;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 244 IMTKYISKPENIKLMMNLIRDKS 266  
DB 245 IMTKYISKPENIKLMMNLIRDKS 267  
RESULT 2  
TS0117  
mo25 homolog [imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 28-Jul-2000  
C:Accession: TS0117  
R.Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, February 2000  
A:Reference number: Z25039  
A:Accession: TS0117  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-329 <SEE>  
A:Cross-references: EMBL:AL157734; PIDN:CAB75774.1; GSPDB:GN000066; SPDB:SPAC1834.06C  
A:Experimental source: strain 972h(-); cosmid c1834  
C:Genetics:  
A:Gene: SPDB:SPAC1834.06C  
A:Map position: 1  
A:Introns: 34/3; 185/3  
C:Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 4.7%; Score 16; DB 2; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 219 SENYVKRQSLKLGE 234  
DB 216 SENYVKRQSLKLGE 231

RESULT 3  
T27129  
hypothetical protein Y53C12A.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T27129  
R.Kershaw, J.; Lennard, N.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z20315  
A:Accession: T27129  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-338 <WIL>  
A:Cross-references: EMBL:Z99277; PIDN:CAB16486.1; GSPDB:GN000020; CESP:Y53C12A.4  
A:Experimental source: clone Y53C12A  
C:Genetics:  
A:Gene: CESP:Y53C12A.4  
A:Map position: 2  
A:Introns: 29/3; 103/3; 136/2; 215/1; 282/3  
C:Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 4.2%; Score 14; DB 2; Length 338;  
Best Local Similarity 100.0%; Pred. No. 2.9e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 105 LRRQIGTRSPVVEY 118  
DB 109 LRRQIGTRSPVVEY 122

RESULT 4  
T16651

hypothetical protein R02E12.2 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
C:Accession: T16651  
R:Leimbach, D.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of *C. elegans* cosmid R02E12.  
A:Reference number: Z18554  
A:Accession: T16651  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-377 <LEI>  
A:Cross-references: EMBL:U53337; NID:G1255833; PID:G1255838; PIDN:AAA96187.1; GSPDB:GN00  
A:Experimental source: strain Bristol N2; clone R02E12  
C:Genetics;  
A:Gene: CESP.R02E12.2  
A:Map position: X  
A:Introns: 37/3; 146/2; 225/1; 315/3  
C:Superfamily: Saccharomycetes hypothetical protein YKL189W

```
Query Match      4.2%; Score 14; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY            105 LRRQIGTRSP TVEY 118  
               |||||  
Db            119 LRRQIGTRSP TVEY 132

RESULT 5  
 S03900  
 gene corso protein precursor - fruit fly (*Drosophila melanogaster*)  
 N;Alternate names: growth-factor receptor tyrosine kinase  
 C;Species: *Drosophila melanogaster*  
 C;Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 24-Sep-1999  
 C;Accession: S03900

R; Sprenger, F.; Stevens, L.M.; Nusslein-Volhard, C.  
Nature 338, 478-483, 1989  
A; Title: The *Drosophila* gene *torso* encodes a putative receptor tyrosine kinase.  
A; Reference number: S03900; MUID:89181943; PMID:2927509  
A; Accession: S03900  
A; Status: not compared with conceptual translation  
A; Molecule type: DNA  
A; Residues: 1-923 <SPR>  
A; Cross-references: GB:X15150; NID:g8712; PIDN:CAA33247.1; PID:S603810  
C; Genetics:  
A; Gene: *tor*

A; Cross-references: FlyBase:FBgn0003733  
A; Map position: 2R 43D3-E7  
C; Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologs  
C; Keywords: ATP; growth factor receptor; kinase-related transforming protein; transmembrane protein  
F; 1-20/Domain: signal sequence #status predicted <SIG>  
F; 21-923/Product: gene torso protein #status predicted <MAT>  
F; 473-880/Domain: protein kinase homology <KIN>  
F; 481-489/Region: protein kinase ATP-binding motif

```
Query Match      2.4%; Score 8; DB 2; Length 923;
Best Local Similarity 100.0%; Pred.No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 75 SSGLLVTL 82  
DQ 449 SSGLLVTL 456

RESULT 6  
S72925 hypothetical protein B2168 Cl.182 - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C/Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 23-Mar-2001  
C/Accession: S72925  
R/Smith, D.R.; Robinson, K.  
submitted to the EMBL Data Library, November 1993

A,Description: Mycobacterium leprae cosmid B2168.  
A,Reference number: S72586  
A,Accession: S72925  
A,Status: preliminary  
A,Molecule type: DNA  
A,Molecule type: DNA  
A,Residues: 1-103 <MI>  
A,Cross-references: EMBL:U00018; NID:G467037; PIDN:AAA17261.1; PID:G467077

```
Query Match      2.1%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred.No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	15 AEIVKIL 21
	↑↑↑↑↑↑↑
Dp	87 AEIVKIL 93

## RESULT 7

conserved hypothetical protein SA0444 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: G89814

R; Kuroda, M.; Ohka, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K. C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A/Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: G89814

A:STATUS: preliminary

A; status: preliminary  
A: Molecule type: DNA

A; Molecule type: DNA  
A: Residues: 1 123 377

A;Residues: 1-122 <KUR>  
A;Cross-references: GB:BA000018; PID:g13700376; PIDN:BA841674.1; GSPDB:GN00149  
A;Experimental source: strain N315

C:Genetics:

```
Query Match      2.1%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels
```

Qy 22 KDNLAIL 28  
Dp 85 KDNLAIL 91

## RESULT. T a

AG2552  
AG2551  
AG2550  
AG2549  
AG2548  
AG2547  
AG2546  
AG2545  
AG2544  
AG2543  
AG2542  
AG2541  
AG2540  
AG2539  
AG2538  
AG2537  
AG2536  
AG2535  
AG2534  
AG2533  
AG2532  
AG2531  
AG2530  
AG2529  
AG2528  
AG2527  
AG2526  
AG2525  
AG2524  
AG2523  
AG2522  
AG2521  
AG2520  
AG2519  
AG2518  
AG2517  
AG2516  
AG2515  
AG2514  
AG2513  
AG2512  
AG2511  
AG2510  
AG2509  
AG2508  
AG2507  
AG2506  
AG2505  
AG2504  
AG2503  
AG2502  
AG2501  
AG2500  
AG2499  
AG2498  
AG2497  
AG2496  
AG2495  
AG2494  
AG2493  
AG2492  
AG2491  
AG2490  
AG2489  
AG2488  
AG2487  
AG2486  
AG2485  
AG2484  
AG2483  
AG2482  
AG2481  
AG2480  
AG2479  
AG2478  
AG2477  
AG2476  
AG2475  
AG2474  
AG2473  
AG2472  
AG2471  
AG2470  
AG2469  
AG2468  
AG2467  
AG2466  
AG2465  
AG2464  
AG2463  
AG2462  
AG2461  
AG2460  
AG2459  
AG2458  
AG2457  
AG2456  
AG2455  
AG2454  
AG2453  
AG2452  
AG2451  
AG2450  
AG2449  
AG2448  
AG2447  
AG2446  
AG2445  
AG2444  
AG2443  
AG2442  
AG2441  
AG2440  
AG2439  
AG2438  
AG2437  
AG2436  
AG2435  
AG2434  
AG2433  
AG2432  
AG2431  
AG2430  
AG2429  
AG2428  
AG2427  
AG2426  
AG2425  
AG2424  
AG2423  
AG2422  
AG2421  
AG2420  
AG2419  
AG2418  
AG2417  
AG2416  
AG2415  
AG2414  
AG2413  
AG2412  
AG2411  
AG2410  
AG2409  
AG2408  
AG2407  
AG2406  
AG2405  
AG2404  
AG2403  
AG2402  
AG2401  
AG2400  
AG2399  
AG2398  
AG2397  
AG2396  
AG2395  
AG2394  
AG2393  
AG2392  
AG2391  
AG2390  
AG2389  
AG2388  
AG2387  
AG2386  
AG2385  
AG2384  
AG2383  
AG2382  
AG2381  
AG2380  
AG2379  
AG2378  
AG2377  
AG2376  
AG2375  
AG2374  
AG2373  
AG2372  
AG2371  
AG2370  
AG2369  
AG2368  
AG2367  
AG2366  
AG2365  
AG2364  
AG2363  
AG2362  
AG2361  
AG2360  
AG2359  
AG2358  
AG2357  
AG2356  
AG2355  
AG2354  
AG2353  
AG2352  
AG2351  
AG2350  
AG2349  
AG2348  
AG2347  
AG2346  
AG2345  
AG2344  
AG2343  
AG2342  
AG2341  
AG2340  
AG2339  
AG2338  
AG2337  
AG2336  
AG2335  
AG2334  
AG2333  
AG2332  
AG2331  
AG2330  
AG2329  
AG2328  
AG2327  
AG2326  
AG2325  
AG2324  
AG2323  
AG2322  
AG2321  
AG2320  
AG2319  
AG2318  
AG2317  
AG2316  
AG2315  
AG2314  
AG2313  
AG2312  
AG2311  
AG2310  
AG2309  
AG2308  
AG2307  
AG2306  
AG2305  
AG2304  
AG2303  
AG2302  
AG2301  
AG2300  
AG2299  
AG2298  
AG2297  
AG2296  
AG2295  
AG2294  
AG2293  
AG2292  
AG2291  
AG2290  
AG2289  
AG2288  
AG2287  
AG2286  
AG2285  
AG2284  
AG2283  
AG2282  
AG2281  
AG2280  
AG2279  
AG2278  
AG2277  
AG2276  
AG2275  
AG2274  
AG2273  
AG2272  
AG2271  
AG2270  
AG2269  
AG2268  
AG2267  
AG2266  
AG2265  
AG2264  
AG2263  
AG2262  
AG2261  
AG2260  
AG2259  
AG2258  
AG2257  
AG2256  
AG2255  
AG2254  
AG2253  
AG2252  
AG2251  
AG2250  
AG2249  
AG2248  
AG2247  
AG2246  
AG2245  
AG2244  
AG2243  
AG2242  
AG2241  
AG2240  
AG2239  
AG2238  
AG2237  
AG2236  
AG2235  
AG2234  
AG2233  
AG2232  
AG2231  
AG2230  
AG2229  
AG2228  
AG2227  
AG2226  
AG2225  
AG2224  
AG2223  
AG2222  
AG2221  
AG2220  
AG2219  
AG2218  
AG2217  
AG2216  
AG2215  
AG2214  
AG2213  
AG2212  
AG2211  
AG2210  
AG2209  
AG2208  
AG2207  
AG2206  
AG2205  
AG2204  
AG2203  
AG2202  
AG2201  
AG2200  
AG2199  
AG2198  
AG2197  
AG2196  
AG2195  
AG2194  
AG2193  
AG2192  
AG2191  
AG2190  
AG2189  
AG2188  
AG2187  
AG2186  
AG2185  
AG2184  
AG2183  
AG2182  
AG2181  
AG2180  
AG2179  
AG2178  
AG2177  
AG2176  
AG2175  
AG2174  
AG2173  
AG2172  
AG2171  
AG2170  
AG2169  
AG2168  
AG2167  
AG2166  
AG2165  
AG2164  
AG2163  
AG2162  
AG2161  
AG2160  
AG2159  
AG2158  
AG2157  
AG2156  
AG2155  
AG2154  
AG2153  
AG2152  
AG2151  
AG2150  
AG2149  
AG2148  
AG2147  
AG2146  
AG2145  
AG2144  
AG2143  
AG2142  
AG2141  
AG2140  
AG2139  
AG2138  
AG2137  
AG2136  
AG2135  
AG2134  
AG2133  
AG2132  
AG2131  
AG2130  
AG2129  
AG2128  
AG2127  
AG2126  
AG2125  
AG2124  
AG2123  
AG2122  
AG2121  
AG2120  
AG2119  
AG2118  
AG2117  
AG2116  
AG2115  
AG2114  
AG2113  
AG2112  
AG2111  
AG2110  
AG2109  
AG2108  
AG2107  
AG2106  
AG2105  
AG2104  
AG2103  
AG2102  
AG2101  
AG2100  
AG2099  
AG2098  
AG2097  
AG2096  
AG2095  
AG2094  
AG2093  
AG2092  
AG2091  
AG2090  
AG2089  
AG2088  
AG2087  
AG2086  
AG2085  
AG2084  
AG2083  
AG2082  
AG2081  
AG2080  
AG2079  
AG2078  
AG2077  
AG2076  
AG2075  
AG2074  
AG2073  
AG2072  
AG2071  
AG2070  
AG2069  
AG2068  
AG2067  
AG2066  
AG2065  
AG2064  
AG2063  
AG2062  
AG2061  
AG2060  
AG2059  
AG2058  
AG2057  
AG2056  
AG2055  
AG2054  
AG2053  
AG2052  
AG2051  
AG2050  
AG2049  
AG2048  
AG2047  
AG2046  
AG2045  
AG2044  
AG2043  
AG2042  
AG2041  
AG2040  
AG2039  
AG2038  
AG2037  
AG2036  
AG2035  
AG2034  
AG2033  
AG2032  
AG2031  
AG2030  
AG2029  
AG2028  
AG2027  
AG2026  
AG2025  
AG2024  
AG2023  
AG2022  
AG2021  
AG2020  
AG2019  
AG2018  
AG2017  
AG2016  
AG2015  
AG2014  
AG2013  
AG2012  
AG2011  
AG2010  
AG2009  
AG2008  
AG2007  
AG2006  
AG2005  
AG2004  
AG2003  
AG2002  
AG2001  
AG2000  
AG1999  
AG1998  
AG1997  
AG1996  
AG1995  
AG1994  
AG1993  
AG1992  
AG1991  
AG1990  
AG1989  
AG1988  
AG1987  
AG1986  
AG1985  
AG1984  
AG1983  
AG1982  
AG1981  
AG1980  
AG1979  
AG1978  
AG1977  
AG1976  
AG1975  
AG1974  
AG1973  
AG1972  
AG1971  
AG1970  
AG1969  
AG1968

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena cylindrica* ATCC 27801  
A;Reference number: AB1807; MUID:21595285; PMID:11759840

A: Accession: AG2552

A;ACCESSION: AG235Z  
A;STATUS: preliminary

A;status: preliminary

A;Molecule type: DNA

A;Residues: 1-134 <KUR>  
A;Cross-references: GB:AP003603; PIDN:BAB77344.1; PID:gl7134787; GSPDB:GNO0182  
A;Experimental source: strain PCC 7120

U: Getting:  
n, experimenter BO

λ:Comp. 37x8014  
C;Genetics!

A;Gene: alr8014

```
Query Match      2.1%; Score 7; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels
```

QY 24 NLAILEK 30  
Db 104 NLAILEK 110

## RESULT 9

hypothetical protein all7360 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al  
 A:Species: Nostoc sp. PCC 7120  
 C:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AH2522  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AH2522  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-134 <GR>  
 A:Cross-references: KB:BA000020; PIDN:BA077118.1; PID:gl7134559; GSPDB:GN00180  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all7360  
 A:Genome: plasmid

Query Match 2.1%; Score 7; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels

Qy	24 NLAILEK 30
Db	104 NLAILEK 110

## RESULT 10

RZ5001 10  
 F71407  
 hypothetical protein - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 A:Variety: columbiana  
 C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: F71407  
 P:Revan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirkl  
 R.; Wedler, H.; Wedler, R.; Weitzensegger, T.; Pohl, T.M.; Terryn, N.; Giel  
 avanagh, T.; Hempel, S.; Kotzer, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
 Nature 391, 485-488, 1998  
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
 erc, C.; Moore, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
 ch, C.; Chlutz, N.  
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
 A:Reference number: A71400; MUID:98121113; PMID:9461215  
 A:Accession: F71407  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-135 <BEV>  
 A:Cross-references: GB:297336; NID:G2244788; PID:G2244809  
 C:Genetics:  
 A:Map position: 4COP9-4G3845  
 C:Superfamily: Arabidopsis thaliana 15.6K hypothetical protein

```
Query Match      2.1%; Score 7; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels
```

Qy	87	QLIDFEG	93
Db	123	OLIDFEG	129

## RESULT 11

RESULTS 11  
AB1605  
B. subtilis comG operon protein 6 homolog lin1379 [imported] - *Listeria innocua* (strain

C;Species: *Listeria innocua*  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AB1605

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecq, F.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Madueno, E.; Maitournam, A.; Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlan A./Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AB1605  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-154 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC96610.1; PID:g1  
A;Experimental source: strain Clp11262  
C;Genetics:  
A;Gene: lin1379

```
Query Match          2.1%; Score 7; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 102 NNILRRQ 108  
pB 99 NNILRRQ 105

## RESULT 12

C64300  
 hypothetical protein MJ0003 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: C64300  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake  
 rson, J.D.; Sadow, P.W.; Kinkness, R.; Kinkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasch  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: C64300  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-156 <BUL>  
 A:Cross-references: GB:U67459; GB:L77117; NID:G2826236; PIDN:AAB97990.1; RID:g1498760;  
 C:Genetics:  
 A:Map position: FOR4911-5381

```
Query Match      2.1%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels
```

Qy 291 VEILLKN 297  
db 34 VEILLKN 40

RESIT.T 13

G95316  
 hypothetical protein SMA0809 [imported] - Sinorhizobium meliloti (strain 1021) magapla  
 C;Species: Sinorhizobium meliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C;Accession: G95316  
 R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bo  
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.  
 . Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli  
 A;Reference number: A95262; MUID:21396509; PMID:11481432  
 A;Accession: G95316  
 A;Status: preliminary  
 A;Molecule type: DNA

A:Residues: 1-171 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AAK65097.1; PID:gl4523533; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Gallbert, F.; Finan, T.W.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
C:Contents: annotation  
C:Genetics:  
A:Gene: SMO809  
A:Genome: plasmid

Query Match 2.1%; Score 7; DB 2; Length 171;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;

QY 62 PTEVAQA 68  
|||||  
Db 61 PTEVAQA 67

RESULT 14  
G70325  
transcription regulator TetR/AcrR family - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: G70325  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: G70325  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-179 <AQF>  
A:Cross-references: GB:AE000682; NID:G2982979; PIDN:AAO6593.1; PID:G2982980; GB:AE00065  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: acrR3

Query Match 2.1%; Score 7; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;

QY 155 LAKIILF 161  
|||||  
Db 100 LAKIILF 106

RESULT 15  
AB2247  
hypothetical protein alr3529 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AB2247  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AB2247  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-191 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BAF75228.1; PID:gl7132662; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:

A:Gene: alr3529

Query Match 2.1%; Score 7; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;

QY 82 LIADLQL 88  
|||||  
Db 44 LIADLQL 50

RESULT 16  
SI9734  
glutathione transferase (EC 2.5.1.18) - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 18-Jun-1999  
A:Accession: SI9734  
R:Chang, L.H.; Fan, J.Y.; Liu, L.F.; Tsai, S.P.; Tam, M.F.  
Biochem. J. 281, 545-551, 1992  
A:Title: Cloning and expression of a chick liver glutathione S-transferase CL 3 subunit  
A:Reference number: SI9734; MUID:92143826; PMID:1339283  
A:Accession: SI9734  
A:Molecule type: mRNA  
A:Residues: 1-229 <CHA>  
A:Cross-references: EMBL:M38219; NID:G211529; PIDN:AAAG2731.1; PID:G211530  
C:Superfamily: glutathione transferase  
C:Keywords: transferase

Query Match 2.1%; Score 7; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;

QY 213 YEKLLQS 219  
|||||  
Db 41 YEKLLQS 47

RESULT 17  
T22075  
hypothetical protein F41E7.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
A:Accession: T22075  
R:Lennard, N.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19509  
A:Accession: T22075  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-241 <WIL>  
A:Cross-references: EMBL:Z68106; PIDN:CAA92125.1; GSPDB:GN00028; CESP:F41E7.2  
A:Experimental source: clone F41E7  
C:Genetics:  
A:Gene: CESP:F41E7.2  
A:Map position: X  
A:Introns: 27/2; 54/1; 126/3; 161/3

Query Match 2.1%; Score 7; DB 2; Length 241;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
Indels 0;

QY 289 PIVEILL 295  
|||||  
Db 140 PIVEILL 146

RESULT 18  
T23866  
hypothetical protein R03C1.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
A:Accession: T23866  
R:Steward, C.





## RESULT 23

H89906  
hypothetical protein SAl156 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: H89906  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: H89906  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-299 <XUR>  
A:Cross-references: GB:BA000018; PID:g13701117; PIDN:BA842412.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SAl156

Query Match 2.1%; Score 7; DB 2; Length 299;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 KLLQSEN 221

|||||

63 KLLQSEN 69

## RESULT 24

T40402  
forkhead nuclear signaling protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 07-Dec-1999  
C:Accession: T40402  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: Z1926  
A:Accession: T40402  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-301 <LYN>  
A:Cross-references: EMBL:AL011261; PIDN:CAA20309.1; GSPDB:GN00067; SPDB:SPBC3H7.13  
A:Experimental source: strain 972h-; cosmid c3H7  
C:Genetics:  
A:Gene: SPDB:SPBC3H7.13  
A:Map position: 2

Query Match 2.1%; Score 7; DB 2; Length 301;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EVVSKSL 46

|||||

Db 176 EVVSKSL 182

## RESULT 25

G71441  
hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: columbia  
C>Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 18-Aug-2000  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
P.; Wedler, E.; Wedler, E.; Wambutt, R.; Weizsaecker, T.; Pohl, T.M.; Terry, N.; Giel  
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: G71441

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-305 &lt;BBV&gt;

A:Cross-references: GB:Z97343; NID:g2245073; PID:e327051; PID:g2245086

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: Saccharomyces hypothetical protein YKL189W

Query Match 2.1%; Score 7; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAHVFK 278

|||||

Db 240 EAHVFK 246

## RESULT 26

E64762  
probable 2,6-dioxo-6-phenylhexa-3-enoate hydrolase (EC 3.7.1.8) - Escherichia coli (str  
C:Species: Escherichia coli  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
C:Accession: E64762  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: E64762  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-309 <BIAT>  
A:Cross-references: GB:AF000142; GB:U00096; NID:g1786542; PIDN:AAC73452.1; PID:g178654  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: mhpc  
C:Superfamily: tropinesterase  
C:Keywords: aromatic hydrocarbon catabolism; hydrolase; PCB biodegradation

Query Match 2.1%; Score 7; DB 1; Length 309;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ENLKMM 259

|||||

Db 189 ENLKMM 195

## RESULT 27

D90679  
hypothetical protein E6S0404 [imported] - Escherichia coli (strain O157:H7, substrain )  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: D90679  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ge  
A:Reference number: A96629; MUID:21156231; PMID:11258796  
A:Accession: D90679  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-309 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA833827.1; PID:g13359861; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: E6S0404  
C:Superfamily: tropinesterase

Query Match 2.1%; Score 7; DB 2; Length 309;  
Best Local Similarity 100.0%; Pred. No. 59;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ENLKLM 259  
 |||||  
 Db 189 ENLKLM 195

RESULT 28  
 H85529  
 hypothetical protein mhpC [imported] - Escherichia coli (strain O157:H7, substrain EDL93)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: H85529  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: H85529  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-309 <S70>  
 A:Cross-references: GB:AE005174; NID:q12513185; PIDN:AAG54700.1; GSPDB:GN00145; UWGP:Z04  
 C:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: mhpC  
 C:Superfamily: tropinesterase

Query Match 2.1%; Score 7; DB 2; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ENLKLM 259  
 |||||  
 Db 189 ENLKLM 195

RESULT 29  
 T17557  
 procyclin homolog A67R - Chlorella virus PBCV-1  
 C:Species: Chlorella virus PBCV-1  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T17557  
 R:Graves, M.V.; Van Etten, J.L.  
 submitted to the EMBL Data Library, May 1999  
 A:Reference number: Z18806  
 A:Accession: T17557  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-309 <GRA>  
 A:Cross-references: EMBL:U42580; NID:q4028896; PIDN:RAC96435.1  
 A:Experimental source: specific host Chlorella strain NC64A  
 C:Genetics:  
 A:Gene: A67R

Query Match 2.1%; Score 7; DB 2; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 TFDIASD 182  
 |||||  
 Db 120 TFDIASD 126

RESULT 30  
 T43132  
 hypothetical protein - Lactococcus lactis plasmid pmRC01  
 C:Species: Lactococcus lactis  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 C:Accession: T43132  
 R:Dougherty, B.A.; Hill, C.; Weidman, J.F.; Richardson, D.R.; Venter, J.C.; Ross, R.P.  
 Mol. Microbiol. 29, 1029-1038, 1998  
 A:Title: Sequence and analysis of the 60 kb conjugative, bacteriocin-producing plasmid p

A:Reference number: Z22314  
 A:Accession: T43132  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-314 <DOU>  
 A:Cross-references: EMBL:AE001272; PIDN:AAC56050.1  
 A:Experimental source: strain DPC3147  
 C:Genetics:  
 A:Genome: plasmid pmRC01  
 A:Note: ORF00061

Query Match 2.1%; Score 7; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 RDPFKYV 172  
 |||||  
 Db 182 RDPFKYV 188

RESULT 31  
 F70034  
 conserved hypothetical protein yvdo - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: F70034  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertu  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cl  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser  
 skeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F70034  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-320 <KUN>  
 A:Cross-references: GB:Z99121; GB:AL009126; NID:G2635827; PIDN:CAB15458.1; PID:G263596  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yvdo

Query Match 2.1%; Score 7; DB 2; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 ADQLID 90  
 |||||  
 Db 103 ADQLID 109

RESULT 32  
 G90291  
 endoglucanase precursor [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: G90291  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Char  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
 arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: G90291

```
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332 <KUR>
A:Cross-references: GB:AE006641; NID:gl3814564; PIDN:AAK41590.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS01354

Query Match
Best Local Similarity 2.1%; Score 7; DB 2; Length 332;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 FDIASDA 183
Db 184 FDIASDA 190

RESULT 33
G90360
endoglucanase precursor [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: G90360
R:She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arratt, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: G90360
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AE006641; NID:gl3815224; PIDN:AAK42142.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS01949

Query Match
Best Local Similarity 2.1%; Score 7; DB 2; Length 334;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 FDIASDA 183
Db 186 FDIASDA 192

RESULT 34
T08491
Probable transposase - Enterobacter aerogenes plasmid R751
C:Species: Enterobacter aerogenes
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08491
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A>Title: Conservation of the genetic switch between replication and transfer genes of In
A:Reference number: Z16434; MUID:97118926; PMID:8954881
A:Accession: T08491
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-334 <THO>
A:Cross-references: EMBL:U67194; NID:gl572520; PIDN:AAK64435.1; PID:gl572540
C:Genetics:
A:Gene: tnPA
A:Genome: plasmid R751

Query Match
Best Local Similarity 2.1%; Score 7; DB 2; Length 334;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 35
T08494
Probable transposase - Enterobacter aerogenes plasmid R751
C:Species: Enterobacter aerogenes
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08494
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A>Title: Conservation of the genetic switch between replication and transfer genes of I
A:Reference number: Z16434; MUID:97118926; PMID:8954881
A:Accession: T08494
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-334 <THO>
A:Cross-references: EMBL:U67194; NID:gl572520; PIDN:AAK64438.1; PID:gl572543
C:Genetics:
A:Gene: tnPA
A:Genome: plasmid R751

Query Match
Best Local Similarity 2.1%; Score 7; DB 2; Length 334;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 36
S04682
ribosomal protein varl - yeast (Candida glabrata) mitochondrion
C:Species: mitochondrion Candida glabrata
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 07-Dec-1999
C:Accession: S04682
R:Ainley, W.M.; Macreadie, I.G.; Butow, R.A.
J. Mol. Biol. 184, 565-576, 1985
A>Title: varl gene on the mitochondrial genome of Torulopsis glabrata.
A:Reference number: S04681; MUID:86011564; PMID:3900417
A:Accession: S04682
A:Molecule type: DNA
A:Residues: 1-336 <AIN>
A:Cross-references: EMBL:X02893
A>Note: the source is designated as Torulopsis glabrata
C:Genetics:
A:Gene: varl
A:Genome: mitochondrion
A:Genetic code: SGC2
C:Superfamily: Saccharomyces cerevisiae ribosomal protein varl
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match
Best Local Similarity 2.1%; Score 7; DB 2; Length 336;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYIS 250
Db 177 IMTKYIS 183

RESULT 37
PEIKL
polyporopepsin (EC 3.4.23.29) - Irpex lacteus
C:Species: Irpex lacteus
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Dec-1993
C:Accession: JU0057
R:Kobayashi, H.; Sekibata, S.; Shibuya, H.; Yoshida, S.; Kusakabe, I.; Murakami, K.
Agric. Biol. Chem. 53, 1927-1933, 1989
A>Title: Cloning and sequence analysis of cDNA for Irpex lacteus aspartic proteinase.
A:Reference number: JU0057
A:Accession: JU0057
A:Molecule type: mRNA
A:Residues: 1-340 <KOB>
A>Note: the amino-terminal 24 residues were sequenced on the isolated proteinase
```

C:Superfamily: pepsin  
 C:Keywords: aspartic proteinase; glycoprotein; hydrolase  
 F:32.212/Active site: Asp #status predicted  
 F:192,238/Binding site: Asp #hydrolyse (Asn) (covalent) #status predicted

Query Match 2.1% Score 7; DB 1; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 IASDAFA 185  
 |||||  
 Db 220 IASDAFA 225

RESULT 38  
 B36987  
 probable transcription regulator [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: B96987  
 R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium CLO  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: B96987  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-344 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK78685.1; PID:G15023588; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC0708

Query Match 2.1% Score 7; DB 2; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EIVKILK 22  
 |||||  
 Db 335 EIVKILK 341

RESULT 39  
 C72746  
 probable carbamoylphosphate synthetase APE0498 - Aeropyrum pernix (strain Kl)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
 C:Accession: C72746  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyz  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: C72746  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-357 <KAW>  
 A:Cross-references: DDBJ:AP000059; NID:95103911; PIDN:BAW79463.1; PID:d1043249; PID:g510  
 A:Experimental source: strain Kl  
 C:Genetics:  
 A:Gene: APE0498

Query Match 2.1% Score 7; DB 2; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 290 IVEILK 296  
 |||||  
 Db 269 IVEILK 275

RESULT 40

AD3559  
 ornithine cyclodeaminase (EC 4.3.1.12) [imported] - Brucella melitensis (strain 16M)  
 C:Species: Brucella melitensis  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
 C:Accession: AD3559  
 R:DelVecchio, V.G.; Kaputraj, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova  
 ; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete;  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melite:  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AD3559  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-359 <KUR>  
 A:Cross-references: GB:AE008918; PIDN:AAJ53639.1; PID:g17984556; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BM810397  
 A:Map position: II  
 C:Superfamily: ornithine cyclodeaminase  
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase

Query Match 2.1% Score 7; DB 2; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 EKLQSE 220  
 |||||  
 Db 347 EKLQSE 353

RESULT 41  
 F82210  
 amino acid ABC transporter, permease protein VCI360 [imported] - Vibrio cholerae (stra:  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: F82210  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: F82210  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-365 <HEI>  
 A:Cross-references: GB:AE004215; GB:AE003852; NID:g9655842; PIDN:AAF94518.1; GSPDB:GN0  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VCI360  
 A:Map position: 1

Query Match 2.1% Score 7; DB 2; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 GLLVTLLI 83  
 |||||  
 Db 157 GLLVTLLI 163

RESULT 42  
 G96668  
 protein FIN19.7 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 01-Mar-2002  
 C:Accession: G96668  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, (

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G96668  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-368 <STO>  
 A:Cross-references: GB:AB005173; NID:G6633811; PIDN:AAF19670.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F1N19.7  
 A:Map position: 1  
 C:Superfamily: Arabidopsis thaliana hypothetical protein T4F9.90

Query Match 2.1%; Score 7; DB 2; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 BEVSKSL 46  
 Db 76 BEVSKSL 82  
 |||||

## RESULT 43

I64223  
 bifunctional endo-1,4-beta-xylanase homolog - Mycoplasma genitalium  
 C:Species: Mycoplasma genitalium  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
 A:Accession: I64223  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhmann, J.; Nguyen, D.; Uutterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346; PMID:7569993  
 A:Accession: I64223  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-372 <TIGR>  
 A:Cross-references: GB:U99699; GB:I43967; NID:G1045903; PID:G1045904; TIGR:MG217  
 A:Experimental source: strain G-37  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 2.1%; Score 7; DB 2; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 LMMNLLR 263  
 Db 324 LMMNLLR 330  
 |||||

## RESULT 44

G70355  
 8-amino-7-oxononanoate synthase (EC 2.3.1.47) - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-May-2000  
 A:Accession: G70355  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666; PMID:9537320  
 A:Accession: G70355  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-373 <AQF>  
 A:Cross-references: GB:AE000699; NID:G2983238; PIDN:AAC06836.1; PID:G2983240; GB:AE00069  
 A:Experimental source: strain VF5  
 C:Genetics:

A:Gene: bioF  
 C:Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology  
 C:Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate  
 F:29-363/Domain: glycine C-acetyltransferase homology <GCA>  
 F:223/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 2.1%; Score 7; DB 2; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
 Db 274 PKLIEFL 280  
 |||||

## RESULT 45

T16059  
 hypothetical protein F13D11.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001  
 A:Accession: T16059  
 R:Fulton, L.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid F13D11.  
 A:Reference number: S69020  
 A:Accession: T16059  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-376 <FUL>  
 A:Cross-references: EMBL:U40939; NID:G1073175; PID:G1072179; PIDN:AAA81703.1; CESP:F131  
 C:Genetics:  
 A:Gene: CESP:F13D11.4  
 A:Introns: 25/2; 43/1; 80/3; 108/2; 137/3; 227/2; 252/3; 304/1; 344/3  
 C:Superfamily: dihydrokaempferol 4-reductase

Query Match 2.1%; Score 7; DB 2; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 VEILLKN 297  
 Db 58 VEILLKN 64  
 |||||

## RESULT 46

T33761  
 hypothetical protein F42A6.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 01-Dec-2000  
 A:Accession: T33761; T32624  
 R:Antoniou, B.; Smith, A.; Gibson, A.  
 submitted to the EMBL Data Library, October 1998  
 A:Description: The sequence of C. elegans cosmid Y5SH10A.  
 A:Reference number: Z21402  
 A:Accession: T33761  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-381 <ANT>  
 A:Cross-references: EMBL:AF100675; PIDN:AAC69002.1; GSPDB:GN000022; CESP:F42A6.1  
 A:Experimental source: strain Bristol N2; clone Y5SH10A  
 R:Du, Z.; Scheet, P.; Andrews, S.  
 submitted to the EMBL Data Library, December 1997  
 A:Description: The sequence of C. elegans cosmid F42A6.  
 A:Reference number: Z21201  
 A:Accession: T32624  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-270 <DUZ>  
 A:Cross-references: EMBL:AF038613; PIDN:AAB92046.1; GSPDB:GN000022; CESP:F42A6.1  
 A:Experimental source: strain Bristol N2; clone F42A6  
 C:Genetics:  
 A:Gene: CESP:F42A6.1  
 A:Map position: 4

A: Introns: 14/1; 73/1; 133/1; 263/3; 323/3  
 C: Superfamily: Caenorhabditis elegans hypothetical protein Y5SH10A.2

Query Match 2.1%; Score 7; DB 2; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 SSFOKER 312  
 |||||  
 Db 317 SSFOKER 323

RESULT 47  
 S34681  
 hypothetical protein YKL189w - yeast (Saccharomyces cerevisiae)  
 C: Species: Saccharomyces cerevisiae  
 C: Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-Apr-2002  
 C: Accession: S34681; S33963; S38021; S38026  
 R: Wieman, S.; Voss, H.; Schwagaer, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues  
 submitted to the EMBL Data Library, July 1993  
 A: Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome X  
 A: Reference number: S34679  
 A: Accession: S34681  
 A: Molecule type: DNA  
 A: Residues: 1-399 <WIE>  
 A: Cross-references: EMBL:X74151; NID:G450365; PIDN:CAA52249.1; PID:G395236  
 A: Experimental source: strain S288C  
 R: Cheret, G.; Mattheakis, L.C.; Sor, F.  
 Yeast 9, 661-667, 1993  
 A: Title: DNA sequence analysis of the YCN2 region of chromosome XI in Saccharomyces cere  
 A: Reference number: S33960; MUID:93348778; PMID:8394042  
 A: Accession: S33963  
 A: Molecule type: DNA  
 A: Residues: 1-399 <CH>  
 A: Cross-references: NID:G69765; NID:G296985; PIDN:CAA49422.1; PID:G296989  
 R: Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grothues, D.; Senses, C.; Stegemann, J.  
 submitted to the Protein Sequence Database, March 1994  
 A: Reference number: S37825  
 A: Accession: S38021  
 A: Molecule type: DNA  
 A: Residues: 1-399 <W12>  
 A: Cross-references: EMBL:Z28189; NID:G486334; PIDN:CAA82032.1; PID:G486335; MIPS:YKL189w  
 A: Experimental source: strain S288C  
 R: Maia e Silva, A.; Bossler, P.; Villela, C.; Fernandes, L.; Soares, H.; Guerreiro, P.; R  
 submitted to the Protein Sequence Database, March 1994  
 A: Reference number: S38024  
 A: Accession: S38026  
 A: Molecule type: DNA  
 A: Residues: 1-399 <MA1>  
 A: Cross-references: EMBL:Z28189; NID:G486334; PIDN:CAA82032.1; PID:G486335; MIPS:YKL189w  
 A: Experimental source: strain S288C  
 C: Genetics:  
 A: Gene: SGD:HYM1  
 A: Cross-references: SGD:S0001672  
 A: Map position: 11L  
 C: Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 2.1%; Score 7; DB 2; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 YVTKRQS 228  
 |||||  
 Db 244 YVTKRQS 250

RESULT 48  
 G95936  
 hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasamid pSymB  
 C: Species: Sinorhizobium meliloti  
 C: Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C: Accession: G95936  
 R: Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A: Title: The complete sequence of the 1,683-kb pSymB megaplasamid from the N2-fixing en  
 A: Reference number: A95842; MUID:21396508; PMID:11481431  
 A: Accession: G95936  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-400 <KUR>  
 A: Cross-references: GB:AL591985; PIDN:CAC49159.1; PID:G15140644; GSPDB:GN00167  
 A: Experimental source: strain 1021, megaplasamid pSymB  
 R: Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubbe  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A: Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaun  
 hebaul, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, I  
 A: Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A: Reference number: A96039; MUID:21368234; PMID:11474104  
 A: Contents: annotation  
 C: Genetics:  
 A: Gene: Smb21255  
 A: Genome: plasmid

Query Match 2.1%; Score 7; DB 2; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 KASEEVS 43  
 |||||  
 Db 140 KASEEVS 146

RESULT 49  
 S73853  
 hypothetical protein P65 - Mycoplasma pneumoniae (strain ATCC 29342)  
 N: Alternate names: hypothetical protein F10\_orf405  
 C: Species: Mycoplasma pneumoniae  
 A: Variety: ATCC 29342  
 C: Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
 C: Accession: S73853; S49068; S49069  
 R: Himmelsreich, R.; Hilbert, H.; Plagens, H.; Pirkil, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A: Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia  
 A: Reference number: S73327; MUID:97105885; PMID:8948633  
 A: Accession: S73853  
 A: Status: nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 1-405 <HIM>  
 A: Cross-references: EMBL:AE000051; GB:U00089; NID:G1674211; PIDN:AAB96175.1; PID:G1674  
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 R: Proft, T.; Herrmann, R.  
 Mol. Microbiol. 13, 337-348, 1994  
 A: Title: Identification and characterization of hitherto unknown Mycoplasma pneumoniae  
 A: Reference number: S49059; MUID:95075318; PMID:7984111  
 A: Accession: S49068  
 A: Status: nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 40-75 <PRO1>  
 A: Cross-references: EMBL:Z32653; NID:G474085; PIDN:CAA83574.1; PID:G474086  
 A: Experimental source: clone F10-2B  
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994  
 A: Accession: S49069  
 A: Status: nucleic acid sequence not shown; translation not shown  
 A: Molecule type: DNA  
 A: Residues: 83-95 <PRO2>  
 A: Cross-references: EMBL:Z32655; NID:G474087; PIDN:CAA83576.1; PID:G474088  
 A: Experimental source: clone F10-2D  
 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994  
 C: Genetics:  
 A: Genetic code: SGC3  
 C: Keywords: duplication  
 F: 57-96/Region: 40-residue repeat <DUP1>  
 F: 122-161/Region: 40-residue repeat <DUP2>

Query Match 2.1% Score 7; DB 2; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 LMOONLLR 263  
 |||||  
 Db 354 LMMNLLR 360

RESULT 50  
 D95975  
 hypothetical outer membrane protein, similar to Wza, OMA family exoF1 [imported] - Sinoz  
 C;Species: Sinozhizobium melliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C;Accession: D95975  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
 A;Reference number: A95842; MUID:21396508; PMID:11481431  
 A;Accession: D95975  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-421 <KUR>  
 A;Cross-references: GB:AL591985; PIDN:CA049468.1; PID:gl15140954; GSPDB:GN00167  
 A;Experimental source: strain 1021, megaplasmid pSymB  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A;Title: The composite genome of the legume symbiont Sinozhizobium melliloti.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: exoF1; SMD20945  
 A;Genome: plasmid

Query Match 2.1% Score 7; DB 2; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 78;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AQLAQEL 73  
 |||||  
 Db 336 AQLAQEL 342

RESULT 51  
 D96719  
 hypothetical protein T6C23.6 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 24-Aug-2001  
 C;Accession: D96719  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 anser, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khavkin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Iuros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: D96719  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-430 <STO>  
 A;Cross-references: GB:AE005173; NID:96665542; PIDN:AAF22911.1; GSPDB:GN00141  
 C;Genetics:  
 A;Gene: T6C23.6  
 A;Map position: 1  
 C;Superfamily: porphobilinogen synthase

Query Match 2.1% Score 7; DB 2; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 261 LLRDKSP 267  
 |||||  
 Db 362 LLRDKSP 368

RESULT 52  
 F97058  
 Fe-S oxidoreductases CAC1286 [imported] - Clostridium acetobutylicum  
 C;Species: Clostridium acetobutylicum  
 C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
 C;Accession: F97058  
 R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee  
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1  
 A;Reference number: A96900; MUID:21359325; PMID:21359325  
 A;Accession: F97058  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-436 <KUR>  
 A;Cross-references: GB:AE001437; PIDN:AAK79257.1; PID:gl15024215; GSPDB:GN00168  
 A;Experimental source: Clostridium acetobutylicum ATCC824  
 C;Genetics:  
 A;Gene: CAC1286  
 C;Superfamily: conserved hypothetical protein b0835

Query Match 2.1% Score 7; DB 2; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LLGELIL 237  
 |||||  
 Db 428 LLGELIL 434

RESULT 53  
 E90563  
 hypothetical protein MYP1 4130 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C;Species: Mycoplasma pulmonis  
 C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C;Accession: E90563  
 R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.  
 Nucleic Acids Res. 29, 2145-2153, 2001  
 A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul  
 A;Reference number: A99512; MUID:21267165; PMID:11353084  
 A;Accession: E90563  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-440 <KUR>  
 A;Cross-references: GB:AL445566; PID:gl4089827; PIDN:CAC13586.1; GSPDB:GN00153  
 A;Experimental source: strain UAB CTIP  
 C;Genetics:  
 A;Gene: MYP1 4130  
 A;Genetic code: SGC3

Query Match 2.1% Score 7; DB 2; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 IVKILKD 23  
 |||||  
 Db 262 IVKILKD 268

RESULT 54  
 JC4348  
 virulence-mediating protein - Vibrio anguillarum  
 C;Species: Vibrio anguillarum



C>Date: 22-Dec-1995 #sequence\_revision 08-Feb-1996 #text\_change 04-Mar-2000  
 C/Accession: J04348  
 R/Milton, D.L.; Norqvist, A.; Wolf-Watz, H.  
 Gene 164, 95-100, 1995  
 A>Title: Sequence of a novel virulence-mediating gene, *virC*, from *Vibrio anguillarum*.  
 A/Reference number: J04347; MUID:96060945; PMID:7590330  
 A/Accession: J04348  
 A/Molecule type: DNA  
 A/Residues: 1-444 <MIL>  
 A/Cross-references: GB:U17054; NID:G576654; PIDN:AAA86985.1; PID:G576656  
 A/Experimental source: VAN20  
 C/Comment: This gene (*virC*) is essential for the virulence of *Vibrio anguillarum*.  
 C/Genetics:  
 A:Gene: *virC*  
 C/Superfamily: *Vibrio anguillarum* virulence-mediating protein  
 C/Keywords: virulence

Query Match 2.1%; Score 7; DB 2; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 AOELYSS 76  
 |||||  
 Db 130 AOELYSS 136

RESULT 55  
 S46175  
 probable resistance protein YBR293w - yeast (*Saccharomyces cerevisiae*)  
 N/Alternate names: hypothetical protein YBR2109  
 C/Species: *Saccharomyces cerevisiae*  
 C/Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 19-Apr-2002  
 C/Accession: S46175  
 R/Fritz, C.; Hollenberg, C.P.; Kirchrath, L.; Ramezani Rad, M.  
 submitted to the Protein Sequence Database, August 1994  
 A/Reference number: S46175  
 A/Accession: S46175  
 A/Molecule type: DNA  
 A/Residues: 1-474 <FRI>  
 A/Cross-references: EMBL:Z36162; NID:G536749; PIDN:CAA85258.1; PID:G536750; GSPDB:GN0000  
 A/Experimental source: strain S288C  
 C/Genetics:  
 A:Gene: MIPS:YBR293w  
 A/Cross-references: SGD:S0000497  
 A/Map position: 2R  
 C/Superfamily: probable resistance protein YBR293w  
 C/Keywords: transmembrane protein  
 F:35-51/Domain: transmembrane #status predicted <TM1>  
 F:98-114/Domain: transmembrane #status predicted <TM2>  
 F:122-138/Domain: transmembrane #status predicted <TM3>  
 F:172-188/Domain: transmembrane #status predicted <TM4>  
 F:201-217/Domain: transmembrane #status predicted <TM5>  
 F:238-254/Domain: transmembrane #status predicted <TM6>  
 F:304-320/Domain: transmembrane #status predicted <TM7>  
 F:381-397/Domain: transmembrane #status predicted <TM8>  
 F:447-463/Domain: transmembrane #status predicted <TM9>

Query Match 2.1%; Score 7; DB 2; Length 474;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ILKONLA 26  
 |||||  
 Db 462 ILKONLA 468

RESULT 56  
 JE0096  
 myocilin - mouse  
 C/Species: *Mus musculus* (house mouse)  
 C/Date: 19-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000  
 C/Accession: JE0096; JE0198  
 R/Iomarev, S.I.; Tamm, E.R.; Chang, B.

Biochem. Biophys. Res. Commun. 245, 887-893, 1998  
 A>Title: Characterization of the mouse *Myoc/Tigr* gene.  
 A/Reference number: JE0096; MUID:98249809; PMID:9588210  
 A/Accession: JE0096  
 A/Molecule type: mRNA  
 A/Residues: 1-490 <TOM>  
 A/Cross-references: GB:AF039869; NID:G3115382; PIDN:AAC40112.1; PID:G3115383  
 R/Takahashi, H.; Noda, S.; Inamura, Y.; Nagasawa, A.; Kubota, R.; Mashima, Y.; Kudoh, S.  
 Biochem. Biophys. Res. Commun. 248, 104-109, 1998  
 A>Title: Mouse myocilin (*Myoc*) gene expression in ocular tissues.  
 A/Reference number: JE0198; MUID:98340858; PMID:9675094  
 A/Accession: JE0198  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-490 <TAK>  
 A/Cross-references: DDBJ:AB013592; NID:G3374583; PIDN:BAA32031.1; PID:G3374584  
 C/Genetics:  
 A:Gene: *Myoc/Tigr*  
 A/Map position: 1

Query Match 2.1%; Score 7; DB 2; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266  
 |||||  
 Db 130 NLLRDKS 136

RESULT 57  
 JC5830  
 myocilin - human  
 C/Species: *Homo sapiens* (man)  
 C/Date: 20-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 17-Mar-1999  
 C/Accession: JC5830  
 R/Kubota, R.; Kudoh, J.; Mashima, Y.; Asakawa, S.; Minoshima, S.; Hejtmancik, J.F.; Og  
 Biochem. Biophys. Res. Commun. 242, 396-400, 1998  
 A>Title: Genomic organization of the human myocilin gene (MYOC) responsible for primary  
 A/Reference number: JC5830; MUID:98113364; PMID:9446806  
 A/Accession: JC5830  
 A/Molecule type: DNA  
 A/Residues: 1-504 <KUB>  
 A/Cross-references: DDBJ:AB006686  
 C/Comment: This cytoskeletal protein is involved in the morphogenesis of the basal body  
 ucoma.  
 C/Genetics:  
 A:Gene: *myoc*  
 A/Introns: 202/1; 244/1

Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 93;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266  
 |||||  
 Db 144 NLLRDKS 150

RESULT 58  
 E82508  
 methyl-accepting chemotaxis protein VCA0031 [imported] - *Vibrio cholerae* (strain N1696)  
 C/Species: *Vibrio cholerae*  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C/Accession: E82508  
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A/Reference number: E82035; MUID:20406833; PMID:10952301  
 A/Accession: E82508  
 A/Status: preliminary  
 A/Molecule type: DNA

```
A;Residues: 1-521 <HEI>
A;Cross-references: GB:AE004347; GB:AE003853; NID:g9657411; PIDN:AAF95945.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype EI Tor
C;Genetics:
A;Gene: VCA0031
A;Map position: 2

Query Match      2.1%; Score 7; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 96; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 78 LLVTLIA 84
    |||||
Db 177 LLVTLIA 183

RESULT 59
D85538
hypothetical protein Z0521 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D85538
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-523 <STO>
A;Cross-references: GB:AE005174; NID:g12513274; PIDN:AAG54768.1; GSPDB:GN00145; UWGP:Z05
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z0521

Query Match      2.1%; Score 7; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 96; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 154 PLAKIIL 160
    |||||
Db 356 PLAKIIL 362

RESULT 60
H90687
hypothetical protein EC80472 [imported] - Escherichia coli (strain O157:H7, substrain R1
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: H90687
R;Hayaishi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90687
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-523 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA833895.1; PID:g13359929; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain R1MD 0509952
C;Genetics:
A;Gene: EC80472

Query Match      2.1%; Score 7; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 96; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 154 PLAKIIL 160
    |||||
Db 356 PLAKIIL 362
```

## RESULT 61

T27190

hypothetical protein Y55D9A.2a - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T27190

R;Wallis, J.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20325

A;Accession: T27190

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-543 &lt;WIL&gt;

A;Cross-references: EMBL:AL032649; PIDN:CAA21702.1; GSPDB:GN00022; CESP:Y55D9A.2a

A;Experimental source: clone Y55D9A

C;Genetics:

A;Gene: CESP:Y55D9A.2a

A;Map position: 4

A;Introns: 1/3; 44/3; 76/1; 160/3; 209/1; 314/1; 541/3

Query Match 2.1%; Score 7; DB 2; Length 543;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 NEKEPPT 63

|||||

Db 518 NEKEPPT 524

## RESULT 62

T27191

hypothetical protein Y55D9A.2b - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T27191

R;Wallis, J.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20325

A;Accession: T27191

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-552 &lt;WIL&gt;

A;Cross-references: EMBL:AL032649; PIDN:CAA21703.1; GSPDB:GN00022; CESP:Y55D9A.2b

A;Experimental source: clone Y55D9A

C;Genetics:

A;Gene: CESP:Y55D9A.2b

A;Map position: 4

A;Introns: 1/3; 44/3; 76/1; 160/3; 209/1; 314/1

Query Match 2.1%; Score 7; DB 2; Length 552;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 NEKEPPT 63

|||||

Db 518 NEKEPPT 524

## RESULT 63

F86244

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C;Accession: F86244

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: B86244  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-554 <STO>  
 A;Cross-references: GB:AE005172; NID:g2252630; PIDN:AAB65493.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1

Query Match 2.1%; Score 7; DB 2; Length 554;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AILEKQD 32  
 |||||  
 Db 161 AILEKQD 167

RESULT 64  
 A84902  
 auxin-regulated protein GH3 homolog At2g46370 - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 18-Jul-2001  
 C;Accession: A84902  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: A84902  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-576 <STO>  
 A;Cross-references: GB:AE002093; NID:g4559380; PIDN:AAD23040.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g46370  
 A;Map position: 2  
 C;Superfamily: soybean auxin-regulated protein GH3

Query Match 2.1%; Score 7; DB 2; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKQ 298  
 |||||  
 Db 34 EILLKQ 40

RESULT 65  
 T24103  
 hypothetical protein R102.7 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T24103  
 R;Berks, M.  
 submitted to the EMBL Data Library, March 1996  
 A;Reference number: Z19841  
 A;Accession: T24103  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-587 <WIL>  
 A;Cross-references: EMBL:Z70309; PIDN:CAA94361.1; GSPDB:GN00022; CESP:R102.7  
 A;Experimental source: clone R102  
 C;Genetics:  
 A;Gene: CESP:R102.7  
 A;Map position: 4  
 A;Introns: 19/2; 111/3; 147/1; 270/2; 407/1; 481/3; 529/2

Query Match 2.1%; Score 7; DB 2; Length 587;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 KKTDKAS 39  
 |||||  
 Db 356 KKTDKAS 362

RESULT 66  
 T44840  
 Probable dTDPglucose 4,6-dehydratase (EC 4.2.1.46) [imported] - *Acinetobacter lwoffii*  
 C;Species: *Acinetobacter lwoffii*  
 C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000  
 C;Accession: T44840  
 R;Nakar, D.; Gutnick, D.L.  
 submitted to the EMBL Data Library, July 1999  
 A;Description: Genomic organization of the wce region of *Acinetobacter lwoffii* RAG-1 r  
 A;Reference number: Z22856  
 A;Accession: T44840  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-624 <NAK>  
 A;Cross-references: EMBL:AJ243431; PIDN:CAB57208.1  
 A;Experimental source: strain RAG-1  
 C;Genetics:  
 A;Gene: wcek  
 C;Superfamily: trsG protein  
 C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 2.1%; Score 7; DB 2; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84  
 |||||  
 Db 56 LLVTLIA 62

RESULT 67  
 B86369  
 hypothetical protein F508.10 - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C;Accession: B86369  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzial;  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon  
 kek, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: B86369  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-731 <STO>  
 A;Cross-references: GB:AE005172; NID:g4056437; PIDN:AAC98010.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1

Query Match 2.1%; Score 7; DB 2; Length 731;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 KVLVADF 201  
 |||||  
 Db 519 KVLVADF 525

RESULT 68  
 B82122

ferrous iron transport protein B VC2077 [imported] - Vibrio cholerae (strain N16961 send  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 24-Aug-2001  
C;Accession: B82122  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.  
L. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: B82122  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-758 <HEI>  
A;Cross-references: GB:AE004281; GB:AE003852; NID:9656616; PIDN:AAF95223.1; GSPDB:GN001  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC2077  
A;Map position: 1  
C;Superfamily: ferrous iron transport protein B; translation elongation factor Tu homolog

Query Match 2.1%; Score 7; DB 2; Length 758;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 LVTLIAD 85  
|||||  
Db 332 LVTLIAD 338

RESULT 69  
T38081  
probable mitochondrial intermediate peptidase precursor - fission yeast (Schizosaccharom  
C;Species: Schizosaccharomyces pombe  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 17-Mar-2003  
C;Accession: T38081  
R;Connor, R.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, April 1996  
A;Reference number: 221767  
A;Accession: T38081  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-762 <CON>  
A;Cross-references: EMBL:Z70690; PIDN:CAA94628.1; GSPDB:GN000666; SPDB:SPACIF3.10C  
A;Experimental source: strain 972h-; cosmid c1F3  
C;Genetics:  
A;Gene: SPDB:SPACIF3.10C  
A;Map position: 1  
C;Superfamily: thimet oligopeptidase  
C;Keywords: mitochondrion

Query Match 2.1%; Score 7; DB 2; Length 762;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 ADLQID 90  
|||||  
Db 321 ADLQID 327

RESULT 70  
AG2375  
WD-40 repeat-protein [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AG2375  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AG2375

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-786 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BA076258.1; PID:gl7133695; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr4559

Query Match 2.1%; Score 7; DB 2; Length 786;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLALIK 30  
|||||  
Db 282 NLALIK 288

RESULT 71  
A97668  
mannosidase AGR\_C\_4665 (AF126472) [imported] - Agrobacterium tumefaciens (strain C58, C  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C;Accession: A97668  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldmar  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: A97668  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-818 <KUR>  
A;Cross-references: GB:AE007869; PIDN:AAK88298.1; PID:gl5157768; GSPDB:GN00169  
C;Genetics:  
A;Gene: AGR\_C\_4665  
A;Map position: circular chromosome

Query Match 2.1%; Score 7; DB 2; Length 818;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 KLIGELI 236  
|||||  
Db 239 KLIGELI 245

RESULT 72  
AF2892  
mannosidase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C;Accession: AF2892  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell  
Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AF2892  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-818 <KUR>  
A;Cross-references: GB:AE008688; PIDN:AAI43556.1; PID:gl7741069; GSPDB:GN00186  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu2575  
A;Map position: circular chromosome

Query Match 2.1%; Score 7; DB 2; Length 818;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLLGELI 236  
|||||  
Db 239 KLLGELI 245

RESULT 73  
A70363  
mannose-1-phosphate guanylttransferase - Aquifex aeolicus  
C/Species: Aquifex aeolicus  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
R/Accession: A70363  
R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A/Reference number: A70300; MUID:98196666; PMID:9537320  
A/Accession: A70363  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-831 <AQF>  
A/Cross-references: GB:AE000704; MID:g2983301; PIDN:AA06893.1; PID:g2983302; GB:AE00065  
A/Experimental source: strain VFS  
C/Genetics:  
A/Gene: mpq

Query Match 2.1%; Score 7; DB 2; Length 831;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0

QY 219 SENYVTK 225  
|||||  
Db 533 SENYVTK 539

RESULT 74  
AD0279  
probable virulence factor YPO2291 [imported] - Yersinia pestis (strain CO92)  
C/Species: Yersinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C/Accession: AD0279  
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Accession: AD0279  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-846 <KUR>  
A/Cross-references: GB:AL590842; PIDN:CAC91096.1; PID:g15980287; GSPDB:GN00175  
C/Genetics:  
A/Gene: YPO2291

Query Match 2.1%; Score 7; DB 2; Length 846;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0

QY 290 IVEILLK 296  
|||||  
Db 564 IVEILLK 570

RESULT 75  
MBY7C  
probable membrane protein YCR037c - yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein YCR524  
C/Species: Saccharomyces cerevisiae  
C/Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 16-Jun-2000  
C/Accession: S12919; S19449  
R/Thierry, A.; Fairhead, C.; Dujon, B.  
Yeast 6, 521-534, 1990

A/Title: The complete sequence of the 8.2 kb segment left of MAT on chromosome III rev  
A/Reference number: S12916; MUID:91181345; PMID:1964349  
A/Accession: S12919  
A/Molecule type: DNA  
A/Residues: 1-923 <THI>  
A/Cross-references: EMBL:X56909; NID:g4489; PIDN:CAA40229.1; PID:g4493  
R/Herbert, C.J.; Jia, Y.; Slonimski, P.P.  
submitted to the Protein Sequence Database, March 1992  
A/Reference number: S19445  
A/Accession: S19449  
A/Molecule type: DNA  
A/Residues: 1-923 <DUJ>  
A/Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42304.1; PID:g1907179; GSPDB:GN  
C/Genetics:  
A/Gene: SGD:PHO87; MIPS:YCR037c  
A/Cross-references: SGD:S0000633; MIPS:YCR037c  
A/Map position: 3R  
C/Superfamily: probable membrane protein YCR037c  
C/Keywords: transmembrane protein  
F/458-479/Domain: transmembrane #status predicted <TM1>  
F/501-518/Domain: transmembrane #status predicted <TM2>  
F/538-554/Domain: transmembrane #status predicted <TM3>  
F/583-603/Domain: transmembrane #status predicted <TM4>  
F/628-644/Domain: transmembrane #status predicted <TM5>  
F/675-694/Domain: transmembrane #status predicted <TM6>  
F/708-732/Domain: transmembrane #status predicted <TM7>  
F/739-759/Domain: transmembrane #status predicted <TM8>  
F/766-784/Domain: transmembrane #status predicted <TM9>  
F/799-836/Domain: transmembrane #status predicted <TM10>  
F/846-868/Domain: transmembrane #status predicted <TM11>  
F/891-919/Domain: transmembrane #status predicted <TM12>

Query Match 2.1%; Score 7; DB 1; Length 923;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0

QY 75 SSGLVLT 81  
|||||  
Db 784 SSGLVLT 790

RESULT 76  
C43274  
N-methyl D-aspartate receptor (NMDR) glutamate-gated ion channels subtype NR2C - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C/Accession: C43274  
R/Monyer, H.; Sprengel, R.; Schoepfer, R.; Herb, A.; Higuchi, M.; Lomeli, H.; Burnashev  
Science 256, 1217-1221, 1992  
A/Title: Heteromeric NMDA receptors: molecular and functional distinction of subtypes.  
A/Reference number: A43274; MUID:92271257; PMID:1350393  
A/Accession: C43274  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: nucleic acid  
A/Residues: 1-962 <MON>  
A/Experimental source: brain  
A/Note: sequence extracted from NCBI backbone (NCBIP:103274)  
C/Superfamily: N-methyl-D-aspartate receptor 2C; glutamate receptor homology  
C/Keywords: transmembrane protein  
F/425-852/Domain: glutamate receptor homology <GRH>

Query Match 2.1%; Score 7; DB 2; Length 962;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0

QY 63 TEVAQL 69  
|||||  
Db 101 TEVAQL 107

RESULT 77  
AB7364  
OmpA-related protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C;Accession: A87364  
R;Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: A87364  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1055 <STO>  
A;Cross-references: GB:AE005673; NID:gl3422195; PIDN:AAK22909.1; GSPDB:GN00148  
C;Genetics:  
A;Gene: CC0925

Query Match 2.1%; Score 7; DB 2; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy 262 LRDKSPN 268  
| | | | |  
Db 285 LRDKSPN 291

RESULT 78

A45761  
Ca2+-transporting ATPase (EC 3.6.3.8) - Plasmodium yoelii  
C;Species: Plasmodium yoelii  
C;Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 19-Apr-2002  
C;Accession: A45761  
R;Murakami, K.; Tanabe, K.; Takada, S.  
J. Cell Sci. 97, 487-495, 1990  
A;Title: Structure of a Plasmodium yoelii gene-encoded protein homologous to the Ca(2+)-  
A;Reference number: A45761; MUID:91161669; PMID:2150071  
A;Accession: A45761  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1115 <MUR>  
A;Cross-references: GB:X55197; NID:gl0097; PIDN:CAA38982.1; PID:gl0098  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: hydrolase  
F;699-866/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 2.1%; Score 7; DB 2; Length 1115;  
Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy 17 IVKILKD 23  
| | | | |  
Db 784 IVKILKD 790

RESULT 79

A72287  
Hypothetical protein TM1182 - Thermotoga maritima (strain MSBB)  
C;Species: Thermotoga maritima  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C;Accession: A72287  
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.

Nature 399, 323-329, 1999  
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A;Reference number: A72200; MUID:99287316; PMID:10360571  
A;Accession: A72287  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1170 <ARN>  
A;Cross-references: GB:AE001774; GB:AE000512; NID:g4981717; PIDN:AAD36257.1; PID:g498173  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Experimental source: strain MSBB  
C;Genetics:

A;Gene: TM1182

C;Superfamily: chromosome segregation protein SMCI

Query Match 2.1%; Score 7; DB 2; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy 170 KYVELST 176  
| | | | |  
Db 310 KYVELST 316

RESULT 80

I49705  
Glutamate receptor channel subunit epsilon 3 - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000  
C;Accession: I49705  
R;Kutsuwada, T.; Kashiwabuchi, N.; Mori, H.; Sakimura, K.; Kushiya, E.; Araki, K.; Megu  
Nature 358, 36-41, 1992  
A;Title: Molecular diversity of the NMDA receptor channel.  
A;Reference number: I49704; MUID:92310584; PMID:1377365  
A;Accession: I49705  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1239 <RES>  
A;Cross-references: GB:D10694; NID:g538239; PIDN:BAA01536.1; PID:g538240  
C;Superfamily: N-methyl-D-aspartate receptor 2C; glutamate receptor homology  
C;Keywords: neurotransmitter receptor  
F;425-852/Domain: glutamate receptor homology <GRH>

Query Match 2.1%; Score 7; DB 2; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy 63 TEVAQL 69  
| | | | |  
Db 101 TEVAQL 107

RESULT 81

B45219  
N-methyl-D-aspartate receptor chain NMDAR2C - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: B45219  
R;Ishii, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.; Kadotani, H.; Yokoi, M.; Akaza  
J. Biol. Chem. 268, 2836-2843, 1993  
A;Title: Molecular characterization of the family of the N-methyl-D-aspartate receptor  
A;Reference number: A45219; MUID:93155102; PMID:8428958  
A;Accession: B45219  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-1250 <ISH>  
A;Experimental source: brain  
A;Note: sequence extracted from NCBI backbone (NCBIP:124263)  
C;Superfamily: N-methyl-D-aspartate receptor 2C; glutamate receptor homology  
F;438-865/Domain: glutamate receptor homology <GRH>

Query Match 2.1%; Score 7; DB 1; Length 1250;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

Qy 63 TEVAQL 69  
| | | | |  
Db 114 TEVAQL 120

RESULT 82

D84727  
Probable RAD50 DNA repair protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 01-Mar-2002

C;Accession: D84727  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Kob, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Niemman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: D84727  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-1292 <STO>  
 A;Cross-references: GB:AE002093; NID:g4263721; PIDN:AAD15407.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g31970  
 A;Map position: 2  
 C;Superfamily: RAD50 protein

Query Match 2.1%; Score 7; DB 2; Length 1292;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKIL 21  
 Db 1058 AEIVKIL 1064

RESULT 83  
 T38842  
 Probable RAS GTPase-activating-like protein - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C;Accession: T38842  
 R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 Submitted to the EMBL Data Library, August 1995  
 A;Reference number: Z21815  
 A;Accession: T38842  
 A;Status: preliminary; translated from GB/EMBL/DD8J  
 A;Molecule type: DNA  
 A;Residues: 1-1489 <GEN>  
 A;Cross-references: EMBL:Z98530; PIDN:CAB11059.1; GSPDB:GN000066; SPDB:SPAC4F8.13c  
 A;Experimental source: strain 972h-; cosmid c4F8  
 C;Genetics:  
 A;Gene: SPDB:SPAC4F8.13c  
 A;Map position: 1

Query Match 2.1%; Score 7; DB 2; Length 1489;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 HKVLVAD 200  
 Db 925 HKVLVAD 931

RESULT 84  
 C71610  
 Probable membrane associated protein PFB0615c - malaria parasite (Plasmodium falciparum)  
 C;Species: Plasmodium falciparum  
 C;Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Jul-2000  
 C;Accession: C71610  
 R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
 ; Perrea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
 Science 282, 1126-1132, 1998  
 A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
 A;Reference number: A71600; MUID:99021743; PMID:9804551  
 A;Accession: C71610  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-2013 <CAR>  
 A;Cross-references: GB:AE001406; GB:AE001362; PIDN:AACT71912.1; PID:g384523  
 A;Experimental source: clone 307  
 C;Genetics:  
 A;Gene: PFB0615c

Query Match 2.1%; Score 7; DB 2; Length 2013;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 IFNNILR 106  
 Db 539 IFNNILR 545

RESULT 85  
 A42771  
 reticulocyte-binding protein 1 - Plasmodium vivax  
 C;Species: Plasmodium vivax  
 C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 28-Apr-1995  
 C;Accession: A42771  
 R;Galinski, M.R.; Medina, C.C.; Ingravalllo, P.; Barnwell, J.W.  
 Cell 69, 1213-1226, 1992  
 A;Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.  
 A;Reference number: A42771; MUID:92315338; PMID:1617731  
 A;Accession: A42771  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-2829 <GAL>  
 A;Experimental source: Belem strain, merozoites  
 A;Note: sequence extracted from NCBI backbone (NCBIN:108114, NCBIP:108115)

Query Match 2.1%; Score 7; DB 2; Length 2829;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234  
 Db 364 SLKLGE 370

RESULT 86  
 S28381  
 utrophin - human  
 N;Alternate names: dystrophin-related protein  
 C;Species: Homo sapiens (man)  
 C;Date: 17-Apr-1993 #sequence\_revision 03-Oct-1995 #text\_change 16-Jul-1999  
 C;Accession: S28381; S28914; S03966  
 R;Tinsley, J.M.  
 submitted to the EMBL Data Library, November 1992  
 A;Reference number: S28381  
 A;Accession: S28381  
 A;Molecule type: mRNA  
 A;Residues: 1-3433 <TIN1>  
 A;Cross-references: EMBL:X69086; NID:g34811; PIDN:CAA48829.1; PID:g34812  
 R;Tinsley, J.M.; Blake, D.J.; Roche, A.; Fairbrother, U.; Riss, J.; Byth, B.C.; Knight, I.  
 Nature 360, 591-593, 1992  
 A;Title: Primary structure of dystrophin-related protein.  
 A;Reference number: S28914; MUID:93096045; PMID:1461283  
 A;Accession: S28914  
 A;Molecule type: mRNA  
 A;Residues: 27-246;2839-3343 <TIN2>  
 A;Cross-references: EMBL:X69086  
 R;Love, D.R.; Hill, D.F.; Dickson, G.; Spurr, N.K.; Byth, B.C.; Marsden, R.F.; Walsh, I.  
 Nature 339, 55-58, 1989  
 A;Title: An autosomal transcript in skeletal muscle with homology to dystrophin.  
 A;Reference number: S03966; MUID:89238543; PMID:2541343  
 A;Accession: S03966  
 A;Molecule type: mRNA  
 A;Residues: 2944-3433 <LOV>  
 A;Cross-references: EMBL:X15488; NID:g30933; PIDN:CAA33515.1; PID:g930062  
 C;Comment: This protein is found primarily at the neuromuscular junctions in adult muscle  
 d regenerating muscle.  
 C;Genetics:  
 A;Gene: GDB:UTRN; DMDL  
 A;Cross-references: GDB:119851; OMIM:128240  
 A;Map position: 6q24-6q24  
 C;Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystro



C;Keywords: actin binding; cytoskeleton; leucine zipper; membrane-associated protein; mu  
F;30-248/Domain: alpha-actinin actin-binding domain homology <ACT>  
F;308-417/Domain: spectrin/dystrophin repeat homology <SP01>  
F;418-526/Domain: spectrin/dystrophin repeat homology <SP02>  
F;528-637/Domain: spectrin/dystrophin repeat homology <SP03>  
F;638-685/Region: hinge  
F;686-796/Domain: spectrin/dystrophin repeat homology <SP04>  
F;804-902/Domain: spectrin/dystrophin repeat homology <SP05>  
F;906-1013/Domain: spectrin/dystrophin repeat homology <SP06>  
F;1015-1122/Domain: spectrin/dystrophin repeat homology <SP07>  
F;1124-1230/Domain: spectrin/dystrophin repeat homology <SP08>  
F;1232-1334/Domain: spectrin/dystrophin repeat homology <SP09>  
F;1339-1450/Domain: spectrin/dystrophin repeat homology <SP10>  
F;1451-1543/Domain: spectrin/dystrophin repeat homology <SP11>  
F;1543-1649/Domain: spectrin/dystrophin repeat homology <SP12>  
F;1651-1755/Domain: spectrin/dystrophin repeat homology <SP13>  
F;1856-1973/Domain: spectrin/dystrophin repeat homology <SP14>  
F;1975-2081/Domain: spectrin/dystrophin repeat homology <SP15>  
F;2083-2185/Domain: spectrin/dystrophin repeat homology <SP16>  
F;2227-2333/Domain: spectrin/dystrophin repeat homology <SP17>  
F;2335-2440/Domain: spectrin/dystrophin repeat homology <SP18>  
F;2442-2556/Domain: spectrin/dystrophin repeat homology <SP19>  
F;2558-2688/Domain: spectrin/dystrophin repeat homology <SP20>  
F;2690-2797/Domain: spectrin/dystrophin repeat homology <SP21>  
F;2798-2869/Region: hinge  
F;2812-2849/Domain: WW repeat homology <WW1>  
F;2837-3117/Region: cysteine-rich  
F;3263-3284/Region: leucine zipper motif  
F;3328-3349/Region: leucine zipper motif

Query Match 2.1%; Score 7; DB 1; Length 3433;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 VKILKDN 24  
|||  
Db 1182 VKILKDN 1188

RESULT 87  
T30851  
lysosomal trafficking regulator, long splice form - mouse  
N;Alternate names: beige protein homolog  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 24-Nov-1999  
C;Accession: T30851  
R;Barbosa, M.D.F.S.; Tchervnev, V.T.; Kingsmore, S.F.  
submitted to the EMBL Data Library, September 1996  
A;Description: Two bg or not two bg? Longest isoform of mouse Lyst (beige) gene.  
A;Reference number: Z20903  
A;Accession: T30851  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-3788 <BAR>  
A;Cross-references: EMBL:U70015; NID:g1813541; PID:g1813542; PIDN:AAC53011.1  
A;Experimental source: strain C57BL/6J  
C;Genetics:  
A;Gene: Lyst  
A;Map position: 1  
C;Keywords: alternative splicing

Query Match 2.1%; Score 7; DB 2; Length 3788;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 SLKLGE 234  
|||  
Db 1204 SLKLGE 1210

RESULT 88  
T13960  
beige protein homolog - rat

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C;Accession: T13960  
R;Mori, M.; Nishikawa, T.; Higuchi, K.; Nishimura, M.  
submitted to the EMBL Data Library, November 1998  
A;Description: Deletion in the beige gene of the beige rat due to recombination between  
A;Reference number: Z17837  
A;Accession: T13960  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-3788 <MOR>  
A;Cross-references: EMBL:AB020019; NID:d1241953; PID:d1035670; PIDN:BAA34688.1  
A;Experimental source: strain DA; spleen  
C;Genetics:  
A;Gene: beige

Query Match 2.1%; Score 7; DB 2; Length 3788;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 SLKLGE 234  
|||  
Db 1203 SLKLGE 1209

RESULT 89  
JP0101  
fibrinogen alpha chain - duck (fragment)  
N;Contains: fibrinopeptide A  
C;Species: Anas platyrhynchos (domestic duck)  
C;Date: 30-Jun-1987 #sequence\_revision 28-Dec-1987 #text\_change 26-Jan-1996  
C;Accession: JP0101  
R;Min, Y.; Ping, Z.; Yaoshi, Z.  
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 28, 31-35, 1985  
A;Title: Purification and primary structures of duck fibrinopeptides A and B.  
A;Reference number: A94238  
A;Accession: JP0101  
A;Molecule type: protein  
A;Residues: 1-15 <MIN>  
C;Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology  
C;Keywords: blood coagulation; plasma; pyroglutamic acid  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 1.8%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 SSFQKE 311  
|||  
Db 5 SSFQKE 10

RESULT 90  
E41606  
homeotic protein Moxc - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 15-Oct-1999  
C;Accession: E41606  
R;Murtha, M.T.; Leckman, J.F.; Ruddle, F.H.  
Proc. Natl. Acad. Sci. U.S.A. 88, 10711-10715, 1991  
A;Title: Detection of homeobox genes in development and evolution.  
A;Reference number: A41606; MUID:92073357; PMID:1720547  
A;Accession: E41606  
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr  
A;Molecule type: DNA  
A;Residues: 1-25 <MU2>  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 1.8%; Score 6; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      247 KYISKP 252
      |||||
Db      3 KYISKP 8

RESULT 91
G45495
beta-defensin-7 - bovine
N:Alternate names: peptide BMBD-7
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Feb-1994 #sequence_revision 22-Apr-1995 #text_change 25-Oct-1996
C:Accession: G45495
R:Selected, M.E.; Tang, Y.Q.; Morris, W.L.; McGuire, P.A.; Novotny, M.J.; Smith, W.; Hens
J. Biol. Chem. 268, 6641-6648, 1993
A:Title: Purification, primary structures, and antibacterial activities of beta-defensin
A:Reference number: A45495; MUID:93203264; PMID:8454635
A:Accession: G45495
A:Molecule type: protein
A:Residues: 1-40 <SEL>
A>Note: sequence modified after extraction from NCBI backbone
C:Keywords: antibacterial; disulfide bond; pyroglutamic acid
F:1-40/Product: beta-defensin-7 #status experimental <MAL>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9-38,16-31,21-39/Disulfide bonds: #status predicted

      Query Match      1.8%; Score 6; DB 2; Length 40;
      Best Local Similarity 100.0%; Pred. No. 95;
      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      106 RRQIGT 111
      |||||
Db      25 RRQIGT 30

RESULT 92
I45495
beta-defensin-9 - bovine
N:Alternate names: peptide BMBD-9
N:Contains: beta-defensin-8
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Feb-1994 #sequence_revision 22-Apr-1995 #text_change 22-Apr-1995
C:Accession: I45495; H45495
R:Selected, M.E.; Tang, Y.Q.; Morris, W.L.; McGuire, P.A.; Novotny, M.J.; Smith, W.; Hens
J. Biol. Chem. 268, 6641-6648, 1993
A:Title: Purification, primary structures, and antibacterial activities of beta-defensin
A:Reference number: A45495; MUID:93203264; PMID:8454635
A:Accession: I45495
A:Molecule type: protein
A:Residues: 1-40 <SEL>
A>Note: sequence modified after extraction from NCBI backbone
A:Accession: H45495
A:Molecule type: protein
A:Residues: 3-40 <SE2>
A>Note: sequence extracted from NCBI backbone (NCBIP:127958)
C:Keywords: pyroglutamic acid
F:1-40/Product: beta-defensin-9 #status experimental <MAL>
F:3-40/Product: beta-defensin-8 #status experimental <MA2>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9-38,16-31,21-39/Disulfide bonds: #status predicted

      Query Match      1.8%; Score 6; DB 2; Length 40;
      Best Local Similarity 100.0%; Pred. No. 95;
      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      106 RRQIGT 111
      |||||
Db      25 RRQIGT 30

RESULT 93
PC4162
toxin-co-regulated protein chain C - Vibrio cholerae (fragment)
C:Species: Vibrio cholerae
```

```
C>Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 08-Oct-1999
C:Accession: PC4162; S37097
R:Ogierman, M.A.; Voss, E.; Meaney, C.; Faast, R.; Attridge, S.R.; Manning, P.A.
Gene 170, 9-16, 1996
A:Title: Comparison of the promoter proximal regions of the toxin-co-regulated tcp gene
A:Reference number: JC4719; MUID:96200848; PMID:8621096
A:Accession: PC4162
A:Molecule type: DNA
A:Residues: 1-46 <OGI>
A:Cross-references: EMBL:X74730; NID:G398396; PIDN:CAA52748.1; PID:G398403
C:Genetics:
A:Gene: tcpC

      Query Match      1.8%; Score 6; DB 2; Length 46;
      Best Local Similarity 100.0%; Pred. No. 1.1e+02;
      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 LKDNLA 26
      |||||
Db      23 LKDNLA 28

RESULT 94
A46257
Dbx homeobox (homeodomain) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Oct-1997
C:Accession: A46257
R:Lu, S.; Bogarad, L.D.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 89, 8053-8057, 1992
A:Title: Expression pattern of a murine homeobox gene, Dbx, displays extreme spatial r
A:Reference number: A46257; MUID:92390387; PMID:11355604
A:Accession: A46257
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-61 <LUL>
A:Experimental source: 13.5 day embryonic telencephalon
A>Note: sequence extracted from NCBI backbone (NCBIP:112797)
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:2-58/Domain: homeobox homology <HOX>

      Query Match      1.8%; Score 6; DB 2; Length 61;
      Best Local Similarity 100.0%; Pred. No. 1.4e+02;
      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      247 KYISKP 252
      |||||
Db      24 KYISKP 29

RESULT 95
T37147
hypothetical protein SCJ9A.11c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37147
R:Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21622
A:Accession: T37147
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-61 <HAR>
A:Cross-references: EMBL:AL109972; PIDN:CAB53272.1; GSPDB:GN00070; SCOEDB:SCJ9A.11c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCJ9A.11c

      Query Match      1.8%; Score 6; DB 2; Length 61;
      Best Local Similarity 100.0%; Pred. No. 1.4e+02;
      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 62 PTEAVA 67  
Db 38 PTEAVA 43

## RESULT 96

E95089  
hypothetical protein SP0772 [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: E95089  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: E95089  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-67 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK74910.1; PID:g14972247; GSPDB:GN00164; TIGR:SP4  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0772

Query Match 1.8%; Score 6; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 LSSFQK 310  
Db 42 LSSFQK 47

## RESULT 97

G71355  
Probable ribosomal protein L29 (rpmC) - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C:Accession: G71355  
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: G71355  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-72 <COL>  
A:Cross-references: GB:AE001202; GB:AE000520; NID:g3322446; PIDN:AAK65182.1; PID:g332246  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0197

Query Match 1.8%; Score 6; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 ILRRQI 109  
Db 45 ILRRQI 50

## RESULT 98

H97956  
hypothetical protein spr0680 [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: H97956  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: H97956  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAK99484.1; PID:g15458268; GSPDB:GN00174  
C:Genetics:  
A:Gene: spr0680

Query Match 1.8%; Score 6; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 LSSFQK 310  
Db 45 LSSFQK 50

## RESULT 99

T42944  
hypothetical protein 30 - ateline herpesvirus 3 (strain 73)  
C:Species: ateline herpesvirus 3  
A:Variety: strain 73  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 19-May-2000  
C:Accession: T42944  
R:Albrecht, J.C.; Fleckenstein, B.  
submitted to the EMBL Data Library, August 1998  
A:Description: Primary structure of the herpesvirus ateles genome.  
A:Reference number: Z22274  
A:Accession: T42944  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-75 <ALB>  
A:Cross-references: EMBL:AF083424; PIDN:AAK955556.1  
A:Experimental source: strain 73  
C:Superfamily: equine herpesvirus 2 hypothetical protein 30

Query Match 1.8%; Score 6; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVSKSL 46  
Db 28 EVSKSL 33

## RESULT 100

A64600  
hypothetical protein HP0641 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: A64600  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.I  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, (c  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: A64600  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-75 <TOM>  
A:Cross-references: GB:AE000578; GB:AE000511; NID:g2313759; PIDN:AAD07716.1; PID:g2313

Query Match 1.8%; Score 6; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 AOELYS 75  
 |||||  
 Db 46 AOELYS 51

## RESULT 101

## AF2836

hypothetical protein Atu2118 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

A:Accession: AF2836  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 i, K.R.P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AF2836

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-80 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAU43108.1; PID:gl7740580; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu2118

A:Map position: circular chromosome

Query Match 1.8%; Score 6; DB 2; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 LQLIDF 91

|||||

Db 71 LQLIDF 76

## RESULT 102

## T10329

hypothetical protein 60 - Orgyia pseudotsugata nuclear polyhedrosis virus  
 C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OPMNPV  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000

A:Accession: T10329

R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.

Virology 229, 381-399, 1997

A:Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis  
 A:Reference number: Z17011; PMID:97271300; PMID:9126251

A:Accession: T10329

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-82 <AHR>

A:Cross-references: EMBL:U75930; NID:g2934903; PID:gl1911306

Query Match 1.8%; Score 6; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 EQFADE 321

|||||

Db 50 EQFADE 55

## RESULT 103

## S75083

hypothetical protein ssl0461 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
 s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75083

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <KAN>

A:Cross-references: EMBL:D90910; GB:AB001339; NID:gl652956; PIDN:BAA17945.1; PID:dl0186  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 1.8%; Score 6; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 71

|||||

Db 32 VAQLAQ 37

## RESULT 104

## A64306

ribosomal protein L31 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

A:Accession: A64306

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,

A:Title: Complete genome sequence of the mechanogenic archaeon, Methanococcus jannaschi

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: A64306

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-87 <BUL>

A:Cross-references: GB:U67463; GB:L77117; NID:gl590846; PIDN:AAB98030.1; PID:gl590847;

C:Genetics:

A:Map position: REV50469-50206

A:Start codon: TTG

C:Superfamily: rat ribosomal protein L31

Query Match 1.8%; Score 6; DB 2; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKI 20

|||||

Db 45 AEIVKI 50

## RESULT 105

## C86200

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

A:Accession: C86200

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonsc

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86200

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-88 <STO>

A;Cross-references: GB:AE005172; NID:98927680; PIDN:AAF82171.1; GSPDB:GN00141  
C;Genetics:  
A;Map position: 1

Query Match 1.8%; Score 6; DB 2; Length 88;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 290 IVEILL 295  
|||||  
Db 43 IVEILL 48

## RESULT 106

E71896

hypothetical protein jhp0717 - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999

C;Accession: E71896

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: E71896

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-88 <ARN>

A;Cross-references: GB:AE001503; GB:AE001439; NID:94155275; PIDN:AAD06300.1; PID:G415528

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp0717

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 88;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 IRDLKK 334  
|||||  
Db 35 IRDLKK 40

## RESULT 107

G90777

hypothetical protein ECs1191 [imported] - Escherichia coli (strain O157:H7, substrain R1

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C;Accession: G90777

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yagunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: G90777

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-89 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA834614.1; PID:gl3360651; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain R1MD 050952

C;Genetics:

A;Gene: ECs1191

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 89;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LAILEX 30  
|||||  
Db 6 LAILEX 11

## RESULT 108

A27056

erythrocyte membrane protein 4.1 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 13-Aug-1999

C;Accession: A27056

R;Ngai, J.; Stack, J.H.; Moon, R.T.; Lazarides, E.

Proc Natl Acad Sci U S A. 84, 4432-4436, 1987

A;Title: Regulated expression of multiple chicken erythroid membrane skeletal protein 4

A;Reference number: A27056; MUID:87260822; PMID:3474611

A;Accession: A27056

A;Molecule type: mRNA

A;Residues: 1-90 <NGA>

A;Cross-references: GB:M16962; NID:9211745; PIDN:AAA48762.1; PID:G211746

C;Superfamily: membrane protein 4.1; protein 4.1 membrane-binding domain homology

C;Keywords: alternative splicing; cytoskeleton; membrane protein

F;1-90/Domain: protein 4.1 membrane-binding domain homology (fragment) <B41>

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 90;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SSGLLV 80  
|||||  
Db 33 SSGLLV 38

## RESULT 109

D64617

hypothetical protein HP0780 - Helicobacter pylori (strain 26695)

C;Species: Helicobacter pylori

C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999

C;Accession: D64617

R;Iomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.L.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I

Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C

A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: D64617

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-90 <TOM>

A;Cross-references: GB:AE000590; GB:AE000511; NID:92313907; PIDN:AAD07833.1; PID:G23139

C;Genetics:

A;Start codon: GTG

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 90;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 IRDLKK 334  
|||||  
Db 37 IRDLKK 42

## RESULT 110

AH1987

hypothetical protein asr1451 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AH1987

R;kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AH1987

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-93 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA873408.1; PID:gl7130796; GSPDB:GN00179

A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: asr1451

Query Match 1.8%; Score 6; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
|||||  
Db 17 VKILKD 22

## RESULT 111

F93332  
Hypothetical protein SM1045 [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: F93332

R: Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bower, K.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C. Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: F93332

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-95 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65224.1; PID:gl4523672; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R: Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, Hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SM1045

A:Genome: plasmid

Query Match 1.8%; Score 6; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLI 83  
|||||  
Db 81 LLVTLI 86

## RESULT 112

AE2631  
Hypothetical protein Atu0448 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AE2631

R: Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E. A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AE2631

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <KUR>

A:Cross-references: GB:AF008688; PIDN:AA141467.1; PID:gl7738792; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu0448  
A:Map position: circular chromosome

Query Match 1.8%; Score 6; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 PIVEIL 294  
|||||  
Db 6 PIVEIL 11

## RESULT 113

G97413  
Hypothetical protein AGR\_C\_795 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C:Accession: G97413

R: Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman  
A: Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: G97413

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86264.1; PID:gi5155372; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C\_795

A:Map position: circular chromosome

Query Match 1.8%; Score 6; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 PIVEIL 294  
|||||  
Db 6 PIVEIL 11

## RESULT 114

AI2262  
Hypothetical protein asl3656 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AI2262

R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AI2262

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-98 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA075355.1; PID:gl7132789; GSPDB:GN00179  
A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asl3656

Query Match 1.8%; Score 6; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 LLKNQP 299  
|||||  
Db 58 LLKNQP 63

## RESULT 115

F82351

gene 3 protein-related protein VC0197 [imported] - Vibrio cholerae (strain N16961 serogp  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82351  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82351  
A:Molecule type: DNA  
A>Status: preliminary  
A:Residues: 1-99 <HEI>  
A:Cross-references: GB:AE004110; GR:AE003852; NID:9654600; PIDN:AAF93373.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0197  
A:Map position: 1

Query Match 1.8%; Score 6; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 LIDFEG 93  
DB 56 LIDFEG 61

RESULT 116  
H70413  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoK1 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 03-Jun-2002  
C:Accession: H70413  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: H70413  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-100 <RQF>  
A:Cross-references: GB:AE000734; NID:92983733; PIDN:RAC07302.1; PID:92983741; GB:AE00069  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: nuoK1  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 1.8%; Score 6; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 LKGYEA 135  
DB 86 LKGYEA 91

RESULT 117  
B44056  
11K protein - canine coronavirus (strain X378)  
N:Alternate names: 6a protein  
C:Species: canine coronavirus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
C:Accession: B44056  
R:Vennema, H.; Rossen, J.W.A.; Wesseling, J.; Horzinek, M.C.; Rottier, P.J.M.  
Virology 191, 134-140, 1992  
A:Title: Genomic organization and expression of the 3' end of the canine and feline ente  
A:Reference number: A44056; MUID:93033103; PMID:1329312  
A:Molecule type: genomic RNA

gene 3 protein-related protein VC0197 [imported] - Vibrio cholerae (strain N16961 serogp  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82351  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82351  
A:Molecule type: DNA  
A>Status: preliminary  
A:Residues: 1-99 <HEI>  
A:Cross-references: GB:AE004110; GR:AE003852; NID:9654600; PIDN:AAF93373.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0197  
A:Map position: 1

Query Match 1.8%; Score 6; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 LIDFEG 93  
DB 56 LIDFEG 61

RESULT 116  
H70413  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoK1 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 03-Jun-2002  
C:Accession: H70413  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: H70413  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-100 <RQF>  
A:Cross-references: GB:AE000734; NID:92983733; PIDN:RAC07302.1; PID:92983741; GB:AE00069  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: nuoK1  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 1.8%; Score 6; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 LKGYEA 135  
DB 86 LKGYEA 91

RESULT 117  
B44056  
11K protein - canine coronavirus (strain X378)  
N:Alternate names: 6a protein  
C:Species: canine coronavirus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
C:Accession: B44056  
R:Vennema, H.; Rossen, J.W.A.; Wesseling, J.; Horzinek, M.C.; Rottier, P.J.M.  
Virology 191, 134-140, 1992  
A:Title: Genomic organization and expression of the 3' end of the canine and feline ente  
A:Reference number: A44056; MUID:93033103; PMID:1329312  
A:Molecule type: genomic RNA

A:Residues: 1-101 <VEN>  
A:Cross-references: GB:X66717; NID:958849; PIDN:CAA47247.1; PID:958851  
C:Superfamily: feline infectious peritonitis virus 11K protein

Query Match 1.8%; Score 6; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLI 83  
DB 96 LLVTLI 101

RESULT 118  
A38354  
carboxypeptidase B (EC 3.4.17.2) precursor - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 30-Dec-1991 #sequence\_revision 30-Dec-1991 #text\_change 09-May-1997  
C:Accession: A38354; B29181  
R:Burgos, F.J.; Salva, M.; Villegas, V.; Soriano, F.; Mendez, E.; Aviles, F.X.  
Biochemistry 30, 4082-4089, 1991  
A:Title: Analysis of the activation process of porcine procarboxypeptidase B and determ  
A:Reference number: A38354; MUID:91208150; PMID:2018774  
A:Accession: A38354  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-103 <BUR>  
R:Aviles, F.X.; Vendrell, J.; Burgos, F.J.; Soriano, F.; Mendez, E.  
Biochem. Biophys. Res. Commun. 130, 97-103, 1985  
A:Title: Sequential homologies between procarboxypeptidases A and B from porcine pancre  
A:Reference number: A29181; MUID:85279427; PMID:4026847  
A:Accession: B29181  
A:Molecule type: protein  
A:Residues: 'SS', 3-13, 'H', 15-23, 'Q', 25-26, 'A', 28-30, 'HXX', 34-38 <AVI>  
C:Superfamily: carboxypeptidase  
C:Keywords: hydrolase; metallo-carboxypeptidase

Query Match 1.8%; Score 6; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 DFLEQN 205  
DB 66 DFLEQN 71

RESULT 119  
A60600  
hypothetical mutant NADH dehydrogenase (ubiquinone) chain 5 - rat mitochondrion (fragme  
C:Species: mitochondrion Rattus norvegicus (Norway rat)  
C:Date: 23-Jul-1998 #sequence\_revision 23-Jul-1998 #text\_change 20-Apr-2000  
C:Accession: A60600  
R:Corral, M.; Paris, B.; Baffet, G.; Tichonicky, L.; Guguen-Guillouzo, C.; Kruh, J.; De  
Exp. Cell Res. 184, 158-166, 1989  
A:Title: Increased level of the mitochondrial ND5 transcript in chemically induced rat  
A:Reference number: A60600; MUID:90005714; PMID:2507335  
A:Accession: A60600  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-107 <COR>  
A:Experimental source: artificial carcinogen induced hepatoma cell line Morris 7288  
C:Genetics:  
A:Gene: ND5  
A:Genome: mitochondrion  
A:Genetic code: SGC1  
C:Keywords: liver; mitochondrion

Query Match 1.8%; Score 6; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLIA 84  
|||||



Db 73 LVTLLA 78

RESULT 120

hypothetical protein APE2596 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000

C:Accession: E72494

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix

A/Reference number: A72450; MUID:93310339; PMID:10382966

A/Accession: E72494

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-108 <KAW>

A/Cross-references: DDBJ:AF000064; NID:g5105945; PIDN:BAR81613.1; PID:d1045399; PID:g510

A/Experimental source: strain K1

C:Genetics:

A/Gene: APE2596

C:Superfamily: Aeropyrum pernix hypothetical protein APE2596

Query Match 1.8%; Score 6; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 YSSGLL 79

Db 28 YSSGLL 33

RESULT 121

AF0091

hypothetical protein YPO0742 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AF0091

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AF0091

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-108 <KUR>

A/Cross-references: GB:AL590842; PIDN:CAC89593.1; PID:g15978825; GSPDB:GN00175

C:Genetics:

A/Gene: YPO0742

Query Match 1.8%; Score 6; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 QSLKLL 232

Db 51 QSLKLL 56

RESULT 122

A83265

conserved hypothetical protein PA3040 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: A83265

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, ; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: A83265

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-109 <STO>

A/Cross-references: GB:AE004729; GB:AE004091; NID:g9949143; PIDN:AAG06428.1; GSPDB:GN0

A/Experimental source: strain PA01

C:Genetics:

A/Gene: PA3040

C:Superfamily: conserved hypothetical protein b2672

Query Match 1.8%; Score 6; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 EKLLQS 219

Db 26 EKLLQS 31

RESULT 123

A89990

hypothetical protein SA1802 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: A89990

R;Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Og ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: A89758; MUID:21311952; PMID:11418146

A/Accession: A89990

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-109 <KUR>

A/Cross-references: GB:BA000018; PID:g13701789; PIDN:BAB43082.1; GSPDB:GN00149

A/Experimental source: strain N315

C:Genetics:

A/Gene: SA1802

Query Match 1.8%; Score 6; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 IRDLKX 334

Db 101 IRDLKX 106

RESULT 124

A24444

hypothetical protein 1 (16S-23S rRNA spacer region) - Chlorella ellipsoidea chloroplast

C:Species: chloroplast Chlorella ellipsoidea

C>Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 18-Jun-1993

C:Accession: A24444

R;Yamada, T.; Shimajiri, M. Nucleic Acids Res. 14, 3827-3839, 1986

A/Title: Peculiar feature of the organization of rRNA genes of the Chlorella chloroplas

A/Reference number: A93622; MUID:86232622; PMID:3714498

A/Accession: A24444

A/Molecule type: DNA

A/Residues: 1-110 <YAM>

A/Note: the authors translated the codon ATT for residue 6 as Asn, CAA for residue 35 a

C:Genetics:

A/Genome: chloroplast

C:Keywords: chloroplast

Query Match 1.8%; Score 6; DB 2; Length 110;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 300 KLIEFL 305

Db 56 KLIEFL 61  
|||||  
A;Cross-references: EMBL:X56770; NID:95589; PIDN:CAA0089.1; PID:95590  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1990

RESULT 125  
T12499  
hypothetical protein DKFZp434K171.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 23-Jul-1999  
C;Accession: T12499  
R;Foustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z17525  
A;Accession: T12499  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-110 <POU>  
A;Cross-references: EMBL:AL080178  
A;Experimental source: adult testis; clone DKFZp434K171  
C;Genetics:  
A;Note: DKFZp434K171.1

Query Match 1.8%; Score 6; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 NILRRQ 108  
|||||  
Db 91 NILRRQ 96

RESULT 126  
S74015  
hypothetical protein c0630 - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 09-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 20-Jun-2000  
C;Accession: S74015  
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.  
Mol. Microbiol. 22, 175-191, 1996  
A;Title: Organizational characteristics and information content of an archaeal genome: 1  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996  
A;Reference number: S73076; MUID:97055432; PMID:88999719  
A;Accession: S74015  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-112 <SEN>  
A;Cross-references: EMBL:Y08256; NID:gl707679; PID:gl707709  
A;Experimental source: strain P2  
C;Superfamily: Sulfolobus solfataricus hypothetical protein c0630

Query Match 1.8%; Score 6; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKIL 21  
|||||  
Db 5 EIVKIL 10

RESULT 127  
S23653  
sensorin A - California sea hare  
C;Species: Aplysia californica (California sea hare)  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
C;Accession: S23653  
R;Brunet, J.F.; Shapiro, E.; Foster, S.A.; Kandel, E.R.; Iino, Y.  
Science 252, 856-859, 1991  
A;Title: Identification of a peptide specific for Aplysia sensory neurons by PCR-based  
A;Reference number: S23653; MUID:91227915; PMID:1840700  
A;Accession: S23653  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-113 <BRU>

Query Match 1.8%; Score 6; DB 2; Length 122;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 IRDLKK 334  
|||||  
Db 57 IRDLKK 62

RESULT 129  
A99504  
hypothetical protein SSO3188 [imported] - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C;Accession: A99504  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: A99504  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-123 <KUR>  
A;Cross-references: GB:AE006641; NID:gl3816623; PIDN:AAK43288.1; GSPDB:GN00155  
C;Genetics:  
A;Gene: SSO3188

Query Match 1.8%; Score 6; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 TVEYIS 120  
|||||  
Db 70 TVEYIS 75

RESULT 130  
A69363  
conserved hypothetical protein AF0905 - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus

Query Match 1.8%; Score 6; DB 2; Length 113;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AILEKQ 31  
|||||  
Db 83 AILEKQ 88

RESULT 128  
C86636  
hypothetical protein yajB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C;Accession: C86636  
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrl  
Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A;Reference number: AB6625; MUID:21235186; PMID:11337471  
A;Accession: C86636  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-122 <STO>  
A;Cross-references: GB:AE005176; PID:gl2722935; PIDN:AAK04189.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: yajB

Query Match 1.8%; Score 6; DB 2; Length 122;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 IRDLKK 334  
|||||  
Db 57 IRDLKK 62

RESULT 129  
A99504  
hypothetical protein SSO3188 [imported] - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C;Accession: A99504  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: A99504  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-123 <KUR>  
A;Cross-references: GB:AE006641; NID:gl3816623; PIDN:AAK43288.1; GSPDB:GN00155  
C;Genetics:  
A;Gene: SSO3188

Query Match 1.8%; Score 6; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 TVEYIS 120  
|||||  
Db 70 TVEYIS 75

RESULT 130  
A69363  
conserved hypothetical protein AF0905 - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 12-Jun-2003  
 C;Accession: A69363  
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 ; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea  
 A;Reference number: A69250; MUID:98049343; PMID:9389475  
 A;Accession: A69363  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-124 <KLE>  
 A;Cross-references: GB:AE001041; GB:AE000782; NID:g2689364; PIDN:AA890335.1; PID:g264969  
 C;Superfamily: uncharacterized conserved protein M01624

Query Match 1.8%; Score 6; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 EGKQDV 97  
 |||||  
 Db 27 EGKQDV 32

RESULT 131  
 F83560  
 Probable type II secretion system protein PA0680 [imported] - Pseudomonas aeruginosa (st  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: F83560  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: F83560  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-124 <STO>  
 A;Cross-references: GB:AE004503; GB:AE004091; NID:g9946553; PIDN:AAG04069.1; GSPDB:GN001  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA0680

Query Match 1.8%; Score 6; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LLRDKS 266  
 |||||  
 Db 42 LLRDKS 47

RESULT 132  
 T06966  
 hypothetical protein ycf35 - Cyanophora paradoxa cyanelle  
 C;Species: cyanelle Cyanophora paradoxa  
 C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 12-Jun-2003  
 C;Accession: T06966  
 R;Stirwalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.  
 submitted to the EMBL Data Library, July 1995  
 A;Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.  
 A;Reference number: Z15840  
 A;Accession: T06966  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-128 <STI>  
 A;Cross-references: EMBL:U00821; NID:gl016083; PIDN:AAA81309.1; PID:gl016222  
 A;Experimental source: strain Fringsheim LB555  
 C;Genetics:

A;Gene: ycf35  
 A;Genome: cyanelle  
 C;Superfamily: uncharacterized conserved protein ycf35  
 C;Keywords: cyanelle

Query Match 1.8%; Score 6; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LIADLQ 87  
 |||||  
 Db 68 LIADLQ 73

RESULT 133  
 A81153  
 type I restriction enzyme-related protein NMB0833 [imported] - Neisseria meningitidis  
 C;Species: Neisseria meningitidis  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C;Accession: A81153  
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.  
 Ricci, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;  
 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A;Reference number: A81000; MUID:20175755; PMID:10710307  
 A;Accession: A81153  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-128 <TET>  
 A;Cross-references: GB:AE002436; GB:AE002098; NID:g7226062; PIDN:AAF41244.1; PID:g72260  
 A;Experimental source: serogroup B, strain MC58  
 C;Genetics:  
 A;Gene: NMB0833

Query Match 1.8%; Score 6; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LIADLQ 87  
 |||||  
 Db 39 LIADLQ 44

RESULT 134  
 H71046  
 hypothetical protein PH1664 - Pyrococcus horikoshii  
 C;Species: Pyrococcus horikoshii  
 C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
 C;Accession: H71046  
 R;Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A;Reference number: A71000; MUID:98344137; PMID:9679194  
 A;Accession: H71046  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-129 <RAW>  
 A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30776.1; PID:g3258093  
 A;Experimental source: strain OT3  
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C;Genetics:  
 A;Gene: PH1664  
 C;Superfamily: Pyrococcus horikoshii hypothetical protein PH1664

Query Match 1.8%; Score 6; DB 2; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGL 78  
 |||||

Db 19 LYSSGL 24

RESULT 135  
H90161  
conserved hypothetical protein [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001  
C:Accession: H90161  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.  
arrett, R.A.; Ragan, M.A.; Seneen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: H90161  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-129 <KUR>  
A:Cross-references: GB:AE006641; NID:gl13813343; PIDN:AAK40551.1; GSPDB:GNC0155  
C:Genetics:  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0505

Query Match 1.8%; Score 6; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 ECIRHE 153  
|||||  
Db 64 ECIRHE 69

RESULT 136  
S48814  
hypothetical protein 2 - turkey herpesvirus  
C:Species: turkey herpesvirus  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S48814  
R:Smith, G.G.; Zelnik, V.V.; Ross, N.N.  
submitted to the EMBL Data Library, October 1994  
A:Description: Gene organization in herpes virus of turkey: identification of a novel OR  
A:Reference number: S48813  
A:Accession: S48814  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-129 <SMI>  
A:Cross-references: EMBL:Z46371; NID:g562786; PIDN:CAA86494.1; PID:g562790

Query Match 1.8%; Score 6; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDA 183  
|||||  
Db 99 DIASDA 104

RESULT 137  
H72722  
hypothetical protein APE0321 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: H72722  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah-  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: H72722  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-130 <KAW>

A:Cross-references: DDBJ:AP000059; NID:g5103911; PIDN:BAA79276.1; PID:g5103960  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0321  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0321

Query Match 1.8%; Score 6; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 IFEDYE 214  
|||||  
Db 67 IFEDYE 72

RESULT 138  
DB4401  
30S ribosomal protein S6E [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: DB4401  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: DB4401  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-131 <STO>  
A:Cross-references: GB:AE004437; NID:gl0581909; PIDN:AAG20576.1; GSPDB:GNC00138  
C:Genetics:  
A:Gene: rps6e  
C:Superfamily: Haloarcula ribosomal protein HS13

Query Match 1.8%; Score 6; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69  
|||||  
Db 103 EAVAQL 108

RESULT 139  
T25924  
hypothetical protein T27E4.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T25924  
R:Bradshaw, H.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid T27E4.  
A:Reference number: Z20111  
A:Accession: T25924  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-131 <BRA>  
A:Cross-references: EMBL:U64837; PIDN:AA04836.1; GSPDB:GNC00023; CESP:T27E4.5  
A:Experimental source: strain Bristol N2; clone T27E4  
C:Genetics:  
A:Gene: CESP:T27E4.5  
A:Map position: 5  
A:Introns: 40/1; 53/2; 98/1; 118/3

Query Match 1.8%; Score 6; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 PLAKII 159  
|||||

Db 3 PLAKII 8

RESULT 140

T30595

sugar transport homolog - Amycolatopsis orientalis

N;Alternate names: PCZA361.28

C;Species: Amycolatopsis orientalis

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C;Accession: T30595

R;Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N

Chem. Biol. 3, 155-162, 1998

A;Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin g

A;Reference number: Z18804

A;Accession: T30595

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-132 <VAN>

A;Cross-references: EMBL:AJ223998

Query Match 1.8%; Score 6; DB 2; Length 132;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 ILFMLL 130

Db 112 ILFMLL 117

RESULT 141

D72054

ribosomal protein S8 CP0113 [imported] - Chlamydomophila pneumoniae (strains CWL029 and AR

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000

C;Accession: D72054; E81612

R;Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A;Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: D72054

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-133 <ARN>

A;Cross-references: GB:AE001647; GB:AE001363; NID:g4376920; PIDN:AA18773.1; PID:g437693

A;Experimental source: strain CWL029

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: B81612

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-133 <REA>

A;Cross-references: GB:AE002173; GB:AE002161; NID:g7189033; PIDN:AAF37996.1; PID:g718904

A;Experimental source: strain AR39, HL cells

C;Genetics:

A;Gene: rsg; CP0113

C;Superfamily: Escherichia coli ribosomal protein S8

Query Match 1.8%; Score 6; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 IVKILK 22

Db 37 IVKILK 42

RESULT 142

G86569

S8 ribosomal protein [imported] - Chlamydomophila pneumoniae (strain J138)

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C;Accession: G86569

R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I

Nucleic Acids Res. 28, 2311-2314, 2000

A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A;Reference number: A86491; MUID:20330349; PMID:10871362

A;Accession: G86569

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-133 <STO>

A;Cross-references: GB:BA000008; NID:g8979006; PIDN:BAA98841.1; GSPDB:GN00142

A;Experimental source: strain J138

C;Genetics:

A;Gene: rsg

C;Superfamily: Escherichia coli ribosomal protein S8

Query Match 1.8%; Score 6; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 IVKILK 22

Db 37 IVKILK 42

RESULT 143

E70334

hypothetical protein aq\_384 - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999

C;Accession: E70334

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; C

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: E70334

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-133 <AQF>

A;Cross-references: GB:AE000687; NID:g2983050; PIDN:AAC06665.1; PID:g2983057; GB:AE0006

A;Experimental source: strain VF5

C;Genetics:

A;Gene: aq\_384

Query Match 1.8%; Score 6; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 KNYLIK 327

Db 79 KNYLIK 84

RESULT 144

S26612

ribosomal protein L27.e, cytosolic - green alga (Pyrobobrys stellata)

C;Species: Pyrobobrys stellata

C;Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 13-Aug-1999

C;Accession: S26612

R;Wolf, A.; Kirsch, M.; Wiessner, W.

submitted to the EMBL Data Library, September 1992

A;Description: Nucleotide sequence of a complementary DNA encoding ribosomal protein L2

A;Reference number: S26612

A;Accession: S26612

A;Molecule type: mRNA

A;Residues: 1-134 <WOL>

A;Cross-references: EMBL:X68202; NID:g18266; PIDN:CAA48289.1; PID:g18267

A;Note: the source is designated as Chlamydomobrys stellata

C;Superfamily: rat ribosomal protein L27

C;Keywords: protein biosynthesis; ribosome

Query Match 1.8%; Score 6; DB 2; Length 134;

Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; VTKRQS 228  
 QY 223 VTKRQS 228  
 Db 53 VTKRQS 58  
 |||||  
 |||||

RESULT 145  
 F72638  
 hypothetical protein APE0542 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C;Accession: F72638  
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yanazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawa, H.; 83-101, 1999  
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1  
 A;Reference number: A72450; MUID:99310339; PMID:10382966  
 A;Accession: F72638  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-134 <RAW>  
 A;Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79510.1; PID:d1043296; PID:g5104188  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APE0542  
 C;Superfamily: Aeropyrum pernix hypothetical protein APE0542

Query Match 1.8%; Score 6; DB 2; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; EFLSSF 308  
 QY 303 EFLSSF 308  
 Db 118 EFLSSF 123  
 |||||  
 |||||

RESULT 146  
 B69103  
 translation initiation factor eIF-2, beta subunit - Methanobacterium thermoautotrophicum  
 C;Species: Methanobacterium thermoautotrophicum  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 08-Oct-1999  
 C;Accession: B69103  
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Ki, S.; Church, G.M.; Daniels, C.G.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function  
 A;Reference number: A69000; MUID:98037514; PMID:9371463  
 A;Accession: B69103  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-135 <MTH>  
 A;Cross-references: GB:AE000932; GB:AE000666; NID:g2622894; PIDN:AA86235.1; PID:g2622894  
 A;Experimental source: strain Delta H  
 C;Genetics:  
 A;Gene: MTH1769  
 C;Superfamily: Methanococcus jannaschii probable translation initiation factor eIF-2 beta subunit

Query Match 1.8%; Score 6; DB 2; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; DYKLL 217  
 QY 212 DYKLL 217  
 Db 3 DYKLL 8  
 |||||  
 |||||

RESULT 147  
 E69842  
 hypothetical protein yJav - Bacillus subtilis  
 C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C;Accession: E69842  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galie iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadachi, Y.; Sato, T.; Scanlon A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama akuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: E69842  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-135 <KUN>  
 A;Cross-references: GB:Z99109; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB12986.1; P A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: yJav  
 C;Superfamily: Bacillus subtilis hypothetical protein yJav

Query Match 1.8%; Score 6; DB 2; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; QNYDTI 209  
 QY 204 QNYDTI 209  
 Db 66 QNYDTI 71  
 |||||  
 |||||

RESULT 148  
 B90443  
 hypothetical protein SSO2684 [imported] - Sulfolobus solfataricus  
 C;Species: Sulfolobus solfataricus  
 C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C;Accession: B90443  
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to Genbank, April 2001  
 A;Description: Sulfolobus solfataricus complete genome.  
 A;Reference number: A99139  
 A;Accession: B90443  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-136 <KUR>  
 A;Cross-references: GB:AE006641; NID:g13816004; PIDN:AAK42801.1; GSPDB:GN00155  
 C;Genetics:  
 A;Gene: SSO2684

Query Match 1.8%; Score 6; DB 2; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; LVTLLA 84  
 QY 79 LVTLLA 84  
 Db 15 LVTLLA 20  
 |||||  
 |||||

RESULT 149  
 T17944  
 hypothetical protein A441L - Chlorella virus PBCV-1  
 C;Species: Chlorella virus PBCV-1  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C;Accession: T17944  
 R;Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806  
 A:Accession: T17944  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-137 <GRA>  
 A:Cross-references: EMBL:U42580; NID:G4028896; PIDN:AAC96809.1  
 A:Experimental source: specific host *Chlorella* strain NC64  
 C:Genetics:  
 A:Note: A441L  
 C:Superfamily: *Chlorella* virus PBCV-1 hypothetical protein A441L

Query Match 1.8%; Score 6; DB 2; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 ILFMILL 130  
 DB 82 ILFMILL 87

# RESULT 150

T07027  
 hypothetical protein (orf in lhca3.St.1 promoter) - potato  
 C:Species: *Solanum tuberosum* (potato)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
 C:Accession: T07027  
 R:Nap, J.P.; van Spanje, M.; Dirkee, W.G.; Baarda, G.; Mlynarova, L.; Loonen, A.; Grondh  
 Plant Mol. Biol. 23: 605-612, 1993  
 A:Title: Activity of the promoter of the lhca3.St.1 gene, encoding the potato apoprotein  
 A:Reference number: S39557; MUID:94033339; PMID:8219093  
 A:Accession: T07027  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-137 <NAP>  
 A:Cross-references: EMBL:S66876; NID:G440964; PIDN:AAB28865.1; PID:G440965

Query Match 1.8%; Score 6; DB 2; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 ILGELI 236  
 DB 117 ILGELI 122

# RESULT 151

E75056  
 probable translation initiation factor aif-2, subunit beta PAB0959 - *Pyrococcus abyssi*  
 C:Species: *Pyrococcus abyssi*  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: E75056  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru  
 A:Reference number: A75001  
 A:Accession: E75056  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-140 <KAW>  
 A:Cross-references: GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB50346.1; PID:G545885  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB0959  
 C:Superfamily: *Methanococcus jannaschii* probable translation initiation factor eIF-2 bet

Query Match 1.8%; Score 6; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYEKLL 217  
 DB 7 DYEKLL 12

# RESULT 152

D71104  
 probable translation initiation factor eIF-2 beta - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: D71104  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch  
 DNA Res. 5: 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: D71104  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-140 <KAW>  
 A:Cross-references: GB:AP000003; NID:G3236130; PIDN:BAA29694.1; PID:G3257011  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0605  
 C:Superfamily: *Methanococcus jannaschii* probable translation initiation factor eIF-2 be

Query Match 1.8%; Score 6; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYEKLL 217  
 DB 7 DYEKLL 12

# RESULT 153

S34667  
 hypothetical protein 140 - *Rhizobium* sp.  
 C:Species: *Rhizobium* sp.  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 15-Jun-2001  
 C:Accession: S34667  
 R:Rochepeau, P.; Feilay, R.; Broughton, W.  
 submitted to the EMBL Data Library, July 1993  
 A:Reference number: S34667  
 A:Accession: S34667  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-140 <ROC>  
 A:Cross-references: EMBL:X74068; NID:G395149; PIDN:CAA52196.1; PID:G581535  
 C:Genetics:  
 A:Start codon: GTG  
 C:Superfamily: *Rhizobium* plasmid pNGR234a protein y4h0

Query Match 1.8%; Score 6; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 RCGIML 146  
 DB 22 RCGIML 27

# RESULT 154

HASH  
 hemoglobin alpha chain - sheep (tentative sequence)  
 C:Species: *Ovis orientalis aries*, *Ovis ammon aries* (domestic sheep)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 03-Mar-2000  
 C:Accession: B90237; A92036; A90574; A02294  
 R:Beale, D.  
 Biochem. J. 103: 129-140, 1967  
 A:Title: A partial amino acid sequence for sheep haemoglobin A.  
 A:Reference number: A90237; MUID:67209244; PMID:6033754  
 A:Accession: B90237  
 A:Molecule type: protein  
 A:Residues: 1-141 <BEA>



A;Experimental source: Soay and Clan breeds  
A;Note: only the composition was determined for the insoluble core region (residues 100-130)  
R;Wilson, J.B.; Brandt, G.; Huismann, T.H.J.  
J. Biol. Chem. 243, 3687-3692, 1968  
A;Title: The structure of sheep hemoglobins. III. Structural studies of the alpha chain  
A;Reference number: A92036; MUID:68313124; PMID:5658545  
A;Contents: composition  
A;Accession: A92036  
A;Molecule type: protein  
A;Residues: 1-103, 'T', 105-123, 'S', 125-141 <WIL>  
A;Experimental source: Rambouillet breed  
R;Huismann, T.H.J.; Dozy, A.M.; Wilson, J.B.; Efremov, G.D.; Vaskov, B.  
Biochim. Biophys. Acta 160, 467-469, 1968  
A;Title: Sheep hemoglobin D, an alpha-chain variant with one apparent amino acid substitution  
A;Reference number: A90574; MUID:69004666; PMID:5690274  
A;Contents: alpha-D allele; composition  
A;Accession: A90574  
A;Molecule type: protein  
A;Residues: 1-14, 'D', 16-141 <HUI>  
C;Superfamily: globin; globin homology  
C;Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier  
F;2-141/Domain: globin homology <GB>  
F;58/Binding site: oxygen (His) [distal axial ligand] #status predicted  
F;87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 1.8%; Score 6; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
Db 104 GLLVTL 109

RESULT 155  
I48552  
orf U5426 - infectious laryngotracheitis virus  
C;Species: infectious laryngotracheitis virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 08-Oct-1999  
C;Accession: I48552  
R;Sakaguchi, M.; Urakawa, T.; Hirayama, Y.; Miki, N.; Yamamoto, M.; Hirai, K.  
Virus Genes 6, 365-378, 1992  
A;Title: Sequence determination and genetic content of an 8.9-kb restriction fragment in  
A;Reference number: A48552; MUID:93118245; PMID:1282282  
A;Accession: I48552  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-142 <SAK>  
A;Cross-references: GB:M80595; NID:g291557; PIDN:AAB59898.1; PID:g291567  
A;Note: sequence extracted from NCBI backbone (NCBIN:121622, NCBIP:121633)

Query Match 1.8%; Score 6; DB 2; Length 142;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
Db 25 SLKLLG 30

RESULT 156  
HHKW41  
heat shock protein 16-41 - Caenorhabditis elegans  
N;Alternate names: heat shock protein 16 2  
C;Species: Caenorhabditis elegans  
C;Date: 25-Feb-1985 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C;Accession: A25199; A38884; A02917  
R;Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.  
J. Biol. Chem. 261, 12006-12015, 1986  
A;Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis  
A;Reference number: A92555; MUID:86304344; PMID:3017958  
A;Accession: A25199  
A;Molecule type: DNA

A;Residues: 1-143 <JON>  
R;Candido, E.P.M.  
submitted to GenBank, November 1985  
A;Reference number: A38884  
A;Accession: A38884  
A;Molecule type: mRNA  
A;Residues: 47-143 <CAN>  
A;Cross-references: GB:X01577; NID:g6758; PIDN:CAA25732.1; PID:g780186  
R;Russnak, R.H.; Jones, D.; Candido, E.P.M.  
Nucleic Acids Res. 11, 3187-3205, 1983  
A;Title: Cloning and analysis of cDNA sequences coding for two 16 kilodalton heat shock  
A;Reference number: A93467; MUID:83220736; PMID:6190129  
A;Accession: A02917  
A;Molecule type: mRNA  
A;Residues: 'KLCSTFQ', 47-143 <RUS>  
A;Cross-references: GB:K01864; NID:g156331; PIDN:AAA28065.1; PID:g156332  
A;Note: the authors translated the codon UUG for residue 46 as Phe  
A;Note: this sequence has been revised in reference A38884  
C;Superfamily: alpha-crystallin  
C;Keywords: heat shock; stress-induced protein

Query Match 1.8%; Score 6; DB 1; Length 143;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPENLK 256  
Db 64 KPENLK 69

RESULT 157  
T15047  
RNA binding protein 3 - wood tobacco  
C;Species: Nicotiana sylvestris (wood tobacco)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: T15047  
R;Moriguchi, K.; Sugita, M.; Sugiyura, M.  
Plant J. 12, 215-221, 1997  
A;Title: Structure and subcellular localization of a small RNA-binding protein from tobacco  
A;Reference number: Z18278; MUID:97408943; PMID:9263462  
A;Accession: T15047  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-144 <MOR>  
A;Cross-references: EMBL:D28862; NID:g2366749; PIDN:BAA22083.1; PID:g2366750  
A;Experimental source: young leaf  
C;Genetics:  
A;Gene: RGP-3  
C;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology  
C;Keywords: RNA binding  
F;39-106/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 1.8%; Score 6; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 DAFATF 187  
Db 56 DAFATF 61

RESULT 158  
T15961  
RNA-binding protein RGP-3 - wood tobacco (fragment)  
C;Species: Nicotiana sylvestris (wood tobacco)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: T15961  
R;Moriguchi, K.; Sugita, M.; Sugiyura, M.  
Plant J. 12, 215-221, 1997  
A;Title: Structure and subcellular localization of a small RNA-binding protein from tobacco  
A;Reference number: Z18278; MUID:97408943; PMID:9263462  
A;Accession: T15961  
A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-144 <MOR>

A:Cross-references: EMBL:D67086; NID:gl009362; PIDN:BAAL1089.1; PID:gl009363

C:Genetics:

A:Introns: 44/1; 67/1; 102/3

C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 1.8%; Score 6; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 182 DAFATF 187

Db 56 DAFATF 61

RESULT 159

AI0008

probable membrane protein YPO0065 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 17-May-2002

C:Accession: AI0008

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11596360

A:Accession: AI0008

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC88931.1; PID:gl5978178; GSPDB:GN00175

C:Genetics:

A:Gene: YPO0065

C:Superfamily: glpE protein

Query Match 1.8%; Score 6; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 HKTQPI 290

Db 93 HKTQPI 98

RESULT 160

B82878

ribosomal protein L9 U551 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000

C:Accession: B82878

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: B82870

A:Accession: B82878

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <GLA>

A:Cross-references: GB:AE002153; GB:AF222894; NID:56899544; PIDN:AAF30964.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: rplJ9; U551

A:Genetic code: SGC3

C:Superfamily: Escherichia coli ribosomal protein L9

Query Match 1.8%; Score 6; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 290 IVEILL 295

Db 125 IVEILL 130

RESULT 161

G69011

hypothetical protein MTH109 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000

C:Accession: G69011

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, I.

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: G69011

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-145 <MTH>

A:Cross-references: GB:AE000801; GB:AE000666; NID:g2621145; PIDN:AAB84615.1; PID:g26211

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH109

C:Superfamily: Methanobacterium thermoautotrophicum hypothetical protein MTH109

Query Match 1.8%; Score 6; DB 2; Length 145;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 DIASDA 183

Db 85 DIASDA 90

RESULT 162

G75031

hypothetical protein PAB0821 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: G75031

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str

A:Reference number: A75001

A:Accession: G75031

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-145 <KAW>

A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50148.1; PID:g54586

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB0821

Query Match 1.8%; Score 6; DB 2; Length 145;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EIVKIL 21

Db 61 EIVKIL 66

RESULT 163

E29010

mer operon OR22 hypothetical protein - Serratia marcescens plasmid pDU1358 (fragment)

C:Species: Serratia marcescens

C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 08-Oct-1999

C:Accession: E29010

R:Griffin, H.G.; Foster, T.J.; Silver, S.; Misra, T.K.

Proc. Natl. Acad. Sci. U.S.A. 84, 3112-3116, 1987

A:Title: Cloning and DNA sequence of the mercuric- and organomercurial-resistance deter

A:Reference number: A94150; MUID:87204087; PMID:3033633

A;Accession: E29010  
A;Molecule type: DNA  
A;Residues: 1-145 <GRA>  
A;Cross-references: GB:M15049; NID:g150627; PIDN:AAA89372.1; PID:g1196707  
C;Genetics:  
A;Genome: plasmid

Query Match 1.8%; Score 6; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 71  
|||||  
Db 11 VAQLAQ 16

RESULT 164  
T10511  
hypothetical protein - Rhodobacter capsulatus  
C;Species: Rhodobacter capsulatus  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-May-2000  
C;Accession: T10511  
R;Barimer, F.W.  
Submitted to the EMBL Data Library, November 1995  
A;Description: Sequence and expression of the pentose-5-phosphate 3-epimerase (cbbE) Cal  
A;Reference number: Z17063  
A;Accession: T10511  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-146 <LAR>  
A;Cross-references: EMBL:U23145; NID:g2564972; PID:g2564977  
A;Experimental source: strain ATCC 11166  
C;Superfamily: Neisseria conserved hypothetical protein NMB1727

Query Match 1.8%; Score 6; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 APOIAL 140  
|||||  
Db 13 APOIAL 18

RESULT 165  
E90104  
40S ribosomal protein S13 [imported] - Guillardia theta nucleomorph  
C;Species: nucleomorph Guillardia theta  
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 15-Jun-2001  
C;Accession: E90104  
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil  
Nature 410, 1091-1096, 2001  
A;Title: The highly reduced genome of an enslaved algal nucleus.  
A;Reference number: A99082; MUID:11323671; PMID:11323671  
A;Accession: E90104  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-147 <DOU>  
A;Cross-references: GB:AJ010592; NID:gl2580664; PIDN:CAC26981.1; GSPDB:GN00151  
C;Genetics:  
A;Map position: 2  
A;Genome: nucleomorph  
C;Superfamily: rat ribosomal protein S13; eubacterial ribosomal protein S15 homology  
C;Keywords: nucleomorph

Query Match 1.8%; Score 6; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILK 22  
|||||  
Db 68 IVKILK 73

RESULT 166  
AD1736  
conserved hypothetical protein lin2433 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C;Accession: AD1736  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Maqueno, E.; Maitournam, A.; M  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant  
A;Title: Comparative genomics of Listeria species  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AD1736  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-147 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC97660.1; PID:g16414955; GSPDB:GN00178  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2433

Query Match 1.8%; Score 6; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 NPAEIV 18  
|||||  
Db 50 NPAEIV 55

RESULT 167  
B71137  
hypothetical protein PH0862 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C;Accession: B71137  
R;Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: B71137  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-148 <KAW>  
A;Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA39956.1; PID:g3257273  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH0862

Query Match 1.8%; Score 6; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKIL 21  
|||||  
Db 64 EIVKIL 69

RESULT 168  
B85641  
hypothetical protein Z1467 [imported] - Escherichia coli (strain O157:H7, substrain EDL  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: B85641  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: B85641  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <STO>  
 A:Cross-references: GB:AE005174; NID:g12514319; PIDN:AG55590.1; GSPDB:GNC0145; UWGP:Z14  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z1467

Query Match 1.8%; Score 6; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 EPTEA 65  
 |||||  
 DB 113 EPTEA 118

## RESULT 169

C90780  
 hypothetical protein ECs1211 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: C90780  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Gasaara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
 A:Reference number: A99629; MUID:21156231; PMID:11259796  
 A:Accession: C90780  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <RAY>  
 A:Cross-references: GB:BA000007; PIDN:BA034634.1; PID:g13360671; GSPDB:GNC00154  
 A:Experimental source: strain O157:H7, substrain R1MD 0509952  
 C:Genetics:  
 A:Gene: ECs1211

Query Match 1.8%; Score 6; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 EPTEA 65  
 |||||  
 DB 113 EPTEA 118

## RESULT 170

AE0928  
 Probable phage tail protein STY3586 [imported] - Salmonella enterica subsp. enterica ser  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AE0928  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parkhill, J.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AE0928  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD09447.1; PID:g16504564; GSPDB:GNC0176  
 C:Genetics:  
 A:Gene: STY3586

Query Match 1.8%; Score 6; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 QLAQEL 73  
 |||||  
 DB 29 QLAQEL 34

## RESULT 171

D90143  
 conserved hypothetical protein [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: D90143  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-NGoc, H.P.; Redder,  
 arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: D90143  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <KUR>  
 A:Cross-references: GB:AE006641; NID:g13813169; PIDN:AAK40403.1; GSPDB:GNC00155  
 C:Genetics:  
 A:Gene: SSO0038

Query Match 1.8%; Score 6; DB 2; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
 |||||  
 DB 93 SLKLLG 98

## RESULT 172

AI2048  
 hypothetical protein all1943 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AI2048  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AI2048  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-149 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA073642.1; PID:g17131033; GSPDB:GNC00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all1943

Query Match 1.8%; Score 6; DB 2; Length 149;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 71  
 |||||  
 DB 95 VAQLAQ 100

## RESULT 173

WVZR2  
 I/K protein - vaccinia virus  
 C:Species: vaccinia virus  
 C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Sep-1999  
 C:Accession: C23768; A26351; A40246  
 R:Weinrich, S.L.; Hruby, D.E  
 Nucleic Acids Res. 14, 3003-3016, 1986

A;Title: A tandemly-oriented late gene cluster within the vaccinia virus genome.

A;Reference number: A93617; MUID:86176781; PMID:3008103

A;Accession: C23768

A;Molecule type: DNA

A;Residues: 1-150 <WEI>

A;Cross-references: GB:X03729; NID:g60836; PIDN:CAA27367.1; PID:g60839

A;Experimental source: strain WR

R;Baldick Jr., C.J.; Moss, B.

Virolology 156, 138-145, 1987

A;Title: Resistance of vaccinia virus to rifampicin conferred by a single nucleotide sub

A;Reference number: A94353; MUID:87122144; PMID:3811229

A;Accession: A26351

A;Molecule type: DNA

A;Residues: 1-150 <BAL>

A;Cross-references: GB:M16556; NID:g335729; PIDN:AAA48304.1; PID:g335730

R;Carpenter, M.S.; DeLange, A.M.

Virolology 188, 233-244, 1992

A;Title: Identification of a temperature-sensitive mutant of vaccinia virus defective in

A;Reference number: A40246; MUID:92230221; PMID:1566576

A;Accession: A40246

A;Molecule type: genomic RNA

A;Residues: 1-150 <CAR>

A;Cross-references: GB:M86531; NID:g335303; PIDN:AAB59802.1; PID:g335304

A;Experimental source: strain WR, mutant tsc63

A;Comment: This protein is probably synthesized after viral DNA replication.

C;Superfamily: vaccinia virus 17K protein

C;Keywords: late protein; transcription regulation

Query Match 1.8%; Score 6; DB 1; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265

Db 121 NLLRDK 126

RESULT 174

F72163

ALL protein - variola minor virus (strain Garcia-1966)

C;Species: variola minor virus

C;Date: 24-Nov-1999 #sequence\_revision 24-Nov-1999 #text\_change 20-Jun-2000

A;Accession: F72163

R;Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Saifonov, P.F.; Massung, R.F.; Lopan

submitted to GenBank, March 1998

A;Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A;Reference number: A72150

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-150 <SHC>

A;Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54704.1; PID:g5830665

A;Experimental source: strain Garcia-1966

C;Genetics:

A;Gene: ALL

C;Superfamily: vaccinia virus 17K protein

Query Match 1.8%; Score 6; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265

Db 121 NLLRDK 126

RESULT 175

C42517

ALL protein - vaccinia virus (strain Copenhagen)

C;Species: vaccinia virus

A;Note: host Homo sapiens (man)

C;Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 16-Feb-1997

C;Accession: C42517

R;Johnson, G.P.

submitted to GenBank, June 1990

A;Reference number: A33172

A;Accession: C42517

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-150 <JOH>

C;Superfamily: vaccinia virus 17K protein

C;Keywords: transcription regulation

Query Match 1.8%; Score 6; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265

Db 121 NLLRDK 126

RESULT 176

B36848

ALL protein - variola virus (strain India-1967)

C;Species: variola virus

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 23-Mar-2001

A;Accession: B36848

R;Blinov, V.M.

submitted to GenBank, November 1992

A;Reference number: A36859

A;Accession: B36848

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-150 <BLI>

A;Cross-references: GB:X69198; NID:g456758; PIDN:CAA49045.1; PID:g297284

C;Superfamily: vaccinia virus 17K protein

C;Keywords: transcription regulation

Query Match 1.8%; Score 6; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265

Db 121 NLLRDK 126

RESULT 177

T28542

hypothetical protein ALL - variola major virus

C;Species: variola major virus

C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000

A;Accession: T28542

R;Massung, R.F.; Eposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubi

Nature 366, 748-751, 1993

A;Title: Potential virulence determinants in terminal regions of variola smallpox virus

A;Reference number: Z20488; MUID:94088747; PMID:8264798

A;Accession: T28542

A;Status: preliminary; translated from GB/EMBL/DBDJB

A;Molecule type: DNA

A;Residues: 1-150 <MAS>

A;Cross-references: EMBL:L22579; NID:g523595; PIDN:AAA60852.1; PID:g439022

A;Experimental source: strain Bangladesh-1975

C;Superfamily: vaccinia virus 17K protein

Query Match 1.8%; Score 6; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265

Db 121 NLLRDK 126

RESULT 178

T37387

VLTF-2, late gene transactivator - vaccinia virus (strain Ankara)  
C;Species: vaccinia virus  
A;Variety: strain Ankara  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000  
C;Accession: T37387  
R;Antoine, G.; Scheiflinger, F.; Palkner, F.G.; Dorner, F.  
submitted to the EMBL Data Library, March 1997  
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain  
A;Reference number: Z20877  
A;Accession: T37387  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-150 <ANT>  
A;Cross-references: EMBL:U94848; PIDN:AA96454.1  
A;Experimental source: strain Ankara  
C;Genetics:  
A;Note: MVA1111  
C;Superfamily: vaccinia virus 17K protein

Query Match 1.8%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDK 265  
Db 121 NLLRDK 126  
|||||

RESULT 179

T22617

hypothetical protein F54B8.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C;Accession: T22617

R;Barlow, K.

submitted to the EMBL Data Library, March 1997

A;Reference number: Z19590

A;Accession: T22617

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-150 &lt;WIL&gt;

A;Cross-references: EMBL:Z93383; PIDN:CAB07622.1; GSPDB:GNO0023; CESP:F54B8.1

A;Experimental source: clone F54B8

C;Genetics:

A;Gene: CESP:F54B8.1

A;Map position: 5

A;Introns: 33/1; 84/2

C;Superfamily: Caenorhabditis elegans hypothetical protein T13C5.3

Query Match 1.8%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 EPPTFA 65  
Db 25 EPPTFA 30  
|||||

RESULT 180

T23641

hypothetical protein M01B2.8 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000

C;Accession: T23641

R;Lloyd, C.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19776

A;Accession: T23641

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-150 &lt;WIL&gt;

A;Cross-references: EMBL:Z83116; PIDN:CAB05566.1; GSPDB:GNO0023; CESP:M01B2.8

A;Experimental source: clone M01B2

C;Genetics:

A;Gene: CESP:M01B2.8

A;Map position: 5

A;Introns: 33/1; 84/2

C;Superfamily: Caenorhabditis elegans hypothetical protein T13C5.3

Query Match 1.8%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 EPPTFA 65  
Db 25 EPPTFA 30  
|||||

RESULT 181

D97188

uncharacterized protein, YBBK B. subtilis ortholog [imported] - Clostridium acetobutyli

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C;Accession: D97188

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: D97188

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-150 &lt;KUR&gt;

A;Cross-references: GB:AE001437; PIDN:AAK80295.1; PID:G15025348; GSPDB:GNO0168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC2339

Query Match

1.8%; Score 6; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 NPAEIV 18  
Db 54 NPAEIV 59  
|||||

RESULT 182

A55209

H transfer determinant A - plasmid R478

C;Species: plasmid R478

C;Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 05-Nov-1999

C;Accession: A55209

R;Whelan, K.F.; Maher, D.; Colleran, E.; Taylor, D.E.

J. Bacteriol. 176, 2242-2251, 1994

A;Title: Genetic and nucleotide sequence analysis of the gene htdA, which regulates cor

A;Reference number: A55209; MUID:94209223; PMID:7908903

A;Accession: A55209

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-150 &lt;WHE&gt;

A;Cross-references: GB:L20341; NID:G442357; PIDN:AA05912.1; PID:G1326033

C;Genetics:

A;Gene: htdA

A;Genome: plasmid

Query Match 1.8%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 229 LKLIGE 234  
Db 139 LKLIGE 144  
|||||

## RESULT 183

AC1758  
 protein gp35 from Bacteriophage A118 homolog lin2608 [imported] - *Listeria innocua* (strain  
 C:Species: *Listeria innocua*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AC1758  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Bolland, J.A.; Voss, H.; Wehlant,  
 A:Title: Comparative Genomics of *Listeria* species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AC1758  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-150 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC97835.1; PID:gl6415145; GSPDB:GN00178  
 A:Experimental source: strain Clp11262  
 C:Genetics:  
 A:Gene: lin2608

Query Match 1.8%; Score 6; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 VADPFL 203  
 |||||  
 Db 106 VADPFL 111

## RESULT 184

AE2193  
 hypothetical protein alr3100 [imported] - *Nostoc* sp. (strain PCC 7120)  
 C:Species: *Nostoc* sp. PCC 7120  
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AE2193  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AE2193  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-151 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BAE74799.1; PID:gl7132194; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr3100  
 C:Superfamily: Haemophilus influenzae conserved hypothetical protein HI0305

Query Match 1.8%; Score 6; DB 2; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 MTKYIS 250  
 |||||  
 Db 1 MTKYIS 6

## RESULT 185

G71820  
 hypothetical protein jhp1337 - *Helicobacter pylori* (strain J99)  
 C:Species: *Helicobacter pylori*  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 29-Sep-1999  
 C:Accession: G71820  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat  
 A:Reference number: A71800; MUID:99120557; PMID:9923682  
 A:Accession: G71820  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <ARN>  
 A:Cross-references: GB:AE001556; GB:AE001439; NID:g4155938; PIDN:AA06913.1; PID:g41559.  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jhp1337  
 C:Superfamily: small protein smpB

Query Match 1.8%; Score 6; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KQSLK 230  
 |||||  
 Db 129 KQSLK 134

## RESULT 186

D64700  
 small protein - *Helicobacter pylori* (strain 26695)  
 C:Species: *Helicobacter pylori*  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 29-Sep-1999  
 C:Accession: D64700  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliak, H.G.; Glodek, A.; McKenn  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C  
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
 A:Reference number: A64520; MUID:97394467; PMID:9252185  
 A:Accession: D64700  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-152 <TOM>  
 A:Cross-references: GB:AE000644; GB:AE000511; NID:g2314609; PIDN:AA08482.1; PID:g23146  
 C:Superfamily: small protein smpB

Query Match 1.8%; Score 6; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KQSLK 230  
 |||||  
 Db 129 KQSLK 134

## RESULT 187

AI3529  
 flbT protein [imported] - *Brucella melitensis* (strain 16M)  
 C:Species: *Brucella melitensis*  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 28-Jul-2003  
 C:Accession: AI3529  
 R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Muje, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
 proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella meliten*  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AI3529  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-152 <KUR>  
 A:Cross-references: GB:AE008918; PIDN:AL53404.1; PID:gl7984299; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEII0163  
 A:Map position: II  
 C:Superfamily: post-transcriptional regulator of flagellin biosynthesis, FlbT type

Query Match 1.8%; Score 6; DB 2; Length 152;



Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 FATPKD 189  
|||||  
Db 90 FATPKD 95

## RESULT 188

AH0438

conserved hypothetical protein YPO3607 [imported] - Yersinia pestis (strain CO92)

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C;Accession: AH0438  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tithball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; et al., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AH0438  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-152 <KUR>  
A;Cross-references: GB:AL590842; PIDN:CAC93076.1; PID:gl59981528; GSPDB:GN00175  
C;Genetics:  
A;Gene: YPO3607

Query Match 1.8%; Score 6; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ILKDNL 25  
|||||  
Db 64 ILKDNL 69

## RESULT 189

T10115

replication-associated protein B - maize streak virus (isolate SP2)

C;Species: maize streak virus  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Nov-2000  
C;Accession: T10115  
R;Tsarnar, M.; Granier, M.; Frutos, R.; Reynaud, B.; Peterschmitt, M. J. Gen. Virol. 79, 3091-3099, 1998  
A;Title: Quasispecies nature of three related maize streak virus isolates obtained through  
A;Reference number: Z16955; MUID:99094636; PMID:9880027  
A;Accession: T10115  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-153 <ISN>  
A;Cross-references: EMBL:AJ225008; NID:g2980718; PIDN:CAAL2319.1; PID:g2980724  
C;Genetics:  
A;Gene: repB  
C;Superfamily: tomato golden mosaic virus AL1 protein  
C;Keywords: DNA replication

Query Match 1.8%; Score 6; DB 2; Length 153;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 PPTPEAV 66  
|||||  
Db 143 PPTPEAV 148

## RESULT 190

S11488

interleukin-2 precursor - sheep

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C;Accession: S11488; S13102; S15517  
R;Goodall, J.C.; Emery, D.C.; Perry, A.C.F.; English, L.S.; Hall, L.

Nucleic Acids Res. 18, 5883, 1990  
A;Title: CDNA cloning of ovine interleukin 2 by PCR.  
A;Reference number: S11488; MUID:91016933; PMID:2216781  
A;Accession: S11488  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-155 <GOO>  
A;Cross-references: EMBL:X53934; NID:gl281; PIDN:CAA37881.1; PID:gl282  
R;Seow, H.F.; Rothel, J.S.; Radford, A.J.; Wood, P.R. Nucleic Acids Res. 18, 7175, 1990

A;Title: The molecular cloning of ovine interleukin 2 gene by the polymerase chain reaction  
A;Reference number: S13102; MUID:91088336; PMID:2263496  
A;Accession: S13102  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-5, L, 7-155 <SBO>  
A;Cross-references: EMBL:X55641; NID:gl810; PIDN:CAA39165.1; PID:gl811  
R;Bujdosó, R.; Williamson, M.L.; Sargan, D.R.; Hein, W.H.; McConnell, I. submitted to the EMBL Data Library, April 1991

A;Reference number: S15517  
A;Accession: S15517  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 21-153 <BUJ>  
A;Cross-references: EMBL:X60148  
C;Superfamily: interleukin-2

Query Match 1.8%; Score 6; DB 2; Length 155;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLXL 257  
|||||  
Db 51 PENLXL 56

## RESULT 191

I45913

interleukin-2 precursor - bovine

C;Species: Bos primigenius taurus (cattle)  
C;Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Jul-1999  
C;Accession: I45913; S21470; S20761  
R;Cerretti, D.P.; McKereghan, K.; Larsen, A.; Cantrell, M.A.; Anderson, D.; Gillis, S.; Proc. Natl. Acad. Sci. U.S.A. 83, 3223-3227, 1986  
A;Title: Cloning, sequence, and expression of bovine interleukin 2.  
A;Reference number: I45913; MUID:86205869; PMID:3517854  
A;Accession: I45913  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-155 <CER>  
A;Cross-references: GB:M12791; NID:gl63204; PIDN:AAA30586.1; PID:gl63205  
R;Anikeeva, N.N.; Vinogradova, T.V.; Votoshin, O.N. submitted to the EMBL Data Library, December 1989

A;Reference number: S21470  
A;Accession: S21470  
A;Molecule type: DNA  
A;Residues: 1-22 <AN2>  
A;Cross-references: EMBL:X17201; NID:g452; PIDN:CAA35062.1; PID:g453  
C;Genetics:  
A;Gene: IL-2  
C;Superfamily: interleukin-2  
C;Keywords: cytokine; glycoprotein; growth factor; immunoregulation; lymphokine; T-cell

Query Match 1.8%; Score 6; DB 2; Length 155;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLXL 257  
|||||  
Db 51 PENLXL 56

## RESULT 192

AC0467  
conserved hypothetical protein YPO3835 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0467  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-156 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC93303.1; PID:gl15981750; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO3835

Query Match 1.8%; Score 6; DB 2; Length 156;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 IRHEPL 155  
Db 83 IRHEPL 88

RESULT 193  
T24304  
hypothetical protein T01G1.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T24304  
R:Kershaw, J.  
Submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19872  
A:Accession: T24304  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-157 <WIL>  
A:Cross-references: EMBL:Z92811; PIDN:CAB07272.1; GSPDB:GN00022; CESP:T01G1.2  
A:Experimental source: clone T01G1  
C:Genetics:  
A:Gene: CESP:T01G1.2  
A:Map position: 4  
A:Introns: 51/3; 107/2; 145/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein T01G1.2

Query Match 1.8%; Score 6; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 ILFMLL 130  
Db 92 ILFMLL 97

RESULT 194  
F87509  
conserved hypothetical protein CC2103 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: F87509  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. n. J.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolch Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: F87509  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-157 <STO>  
A:Cross-references: GB:AE005673; NID:gl3423588; PIDN:AAK24074.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2103

Query Match 1.8%; Score 6; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDA 183  
Db 3 DIASDA 8

RESULT 195  
D90045  
hypothetical protein SA2223 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: D90045  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: D90045  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-157 <KUR>  
A:Cross-references: GB:BA000018; PID:gl3702384; PIDN:BAB43525.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA2223

Query Match 1.8%; Score 6; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 IADLQL 88  
Db 43 IADLQL 48

RESULT 196  
AG1312  
thioredoxin homolog lmo1903 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1312  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker. ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A:Title: Comparative genomics of Listeria species.  
A:Reference number: AS1077; MUID:21537279; PMID:11679669  
A:Accession: AG1312  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-157 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC99981.1; PID:gl6411356; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1903

Query Match 1.8%; Score 6; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 TDKASE 40  
|||||

Db 97 TDKASE 102

RESULT 197  
AG1684  
Choredoxin homolog lin2017 [imported] - *Listeria innocua* (strain Clip11262)  
C:Species: *Listeria innocua*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1684  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative Genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1684  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-157 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC97247.1; PID:gl6414518; GSPDB:GN00178  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin2017

Query Match 1.8%; Score 6; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 TDKASE 40  
Db 97 TDKASE 102

RESULT 198  
D70438  
hypotheoretical protein aq\_1599 - *Aquifex aeolicus*  
C:Species: *Aquifex aeolicus*  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: D70438  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A:Reference number: A70300; MUID:98156666; PMID:9537320  
A:Accession: D70438  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-158 <AQF>  
A:Cross-references: GB:AE000747; NID:g2983944; PIDN:AAC07504.1; PID:g2983956; GB:AE00065  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_1599

Query Match 1.8%; Score 6; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPNLX 256  
Db 97 KPNLX 102

RESULT 199  
C65113  
YnfC protein - *Escherichia coli* (strain K-12)  
C:Species: *Escherichia coli*  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: C65113  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A6720; MUID:97426617; PMID:9278503  
A:Accession: C65113  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-159 <BLAT>  
A:Cross-references: GB:AE000401; GB:U00096; NID:gl789607; PIDN:AAC76249.1; PID:gl789611  
A:Experimental source: strain K-12, substrain MGL655  
C:Genetics:  
A:Gene: ynfC

Query Match 1.8%; Score 6; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 IASDAF 184  
Db 92 IASDAF 97

RESULT 200  
D72226  
conserved hypothetical protein - *Thermotoga maritima* (strain MSB8)  
C:Species: *Thermotoga maritima*  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: D72226  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D. C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: D72226  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-160 <ARN>  
A:Cross-references: GB:AE001808; GB:AE000512; NID:g4982233; PIDN:AAD36745.1; PID:g49822  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM1678

Query Match 1.8%; Score 6; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKMPL 6  
Db 126 MKKMPL 131

RESULT 201  
S73630  
ribosomal protein L10 - *Mycoplasma pneumoniae* (strain ATCC 29342)  
N:Alternate names: hypothetical protein G12\_orf161  
C:Species: *Mycoplasma pneumoniae*  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C:Accession: S73630  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*.  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73630  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-161 <HIM>  
A:Cross-references: EMBL:AE000028; GB:U00089; NID:gl673972; PIDN:AAB95952.1; PID:gl6739  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Gene: rplJ  
A:Genetic code: SGC3  
C:Superfamily: *Escherichia coli* ribosomal protein L10  
C:Keywords: protein biosynthesis; ribosome

```

Query Match          1.8%; Score 6; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107
    |||||
Db 56 NNILRR 61

RESULT 202
AI2638
phaE protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AI2638
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
  erage, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
  ; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
  star, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI2638
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <KUR>
A:Cross-references: GB:AE008688; PIDN:AA141527.1; PID:g17738857; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: phaE
A:Map position: circular chromosome

Query Match          1.8%; Score 6; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKILF 161
    |||||
Db 78 AKILF 83

RESULT 203
I64239
ribosomal protein L10 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 07-Dec-1999
C:Accession: I64239
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
  M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.;
  , C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:7569993
A:Accession: I64239
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-162 <TIGR>
A:Cross-references: GB:U39720; GB:L43967; NID:g1046063; PID:g1046069; TIGR:MG361
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: Escherichia coli ribosomal protein L10

Query Match          1.8%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107
    |||||
Db 57 NNILRR 62

```

```

RESULT 204
B97421
probable poly(3-hydroxyalkanoate) polymerase [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: B97421
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
  A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
  Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
  erum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK86323.1; PID:g15155441; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_998
A:Map position: circular chromosome

Query Match          1.8%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKILF 161
    |||||
Db 79 AKILF 84

RESULT 205
E75100
hypothetical protein PAB1584 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: E75100
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str
  ucture
A:Reference number: A75001
A:Accession: E75100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-164 <KAW>
A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50106.1; PID:g54586:
  6
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1584

Query Match          1.8%; Score 6; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 VTLIAD 85
    |||||
Db 107 VTLIAD 112

RESULT 206
G72486
hypothetical protein APE2535 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: G72486
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takal
  awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
  DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
  us
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: G72486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <KAW>

```

A;Cross-references: DBJ:AP000064; MID:G5105945; PIDN:BAA81551.1; PID:G5106240  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APE2535  
 C;Superfamily: Pyrococcus abyssi hypothetical protein PAB0162

Query Match 1.8%; Score 6; DB 2; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
 |||||  
 Db 35 GLLVTL 40

RESULT 207  
 AD2596  
 biopolymer transport protein exbD [imported] - Agrobacterium tumefaciens (strain C58, Du  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C;Accession: AD2596  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AD2596  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-167 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL41186.1; PID:G17738486; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: exbD  
 A;Map position: circular chromosome  
 C;Superfamily: tolR protein

Query Match 1.8%; Score 6; DB 2; Length 167;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MNLLRD 264  
 |||||  
 Db 120 MNLLRD 125

RESULT 208  
 F97378  
 biopolymer transport exbD protein AGR\_C\_265 [imported] - Agrobacterium tumefaciens (stra  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C;Accession: F97378  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: F97378  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-167 <KUR>  
 A;Cross-references: GB:AE007869; PIDN:AAK85983.1; PID:G15155044; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C\_265  
 A;Map position: circular chromosome  
 C;Superfamily: tolR protein

Query Match 1.8%; Score 6; DB 2; Length 167;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MNLLRD 264  
 |||||  
 Db 120 MNLLRD 125

RESULT 209  
 S42070  
 early chorion protein - silkworm  
 C;Species: Bombyx mori (silkworm)  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Aug-1999  
 C;Accession: S42070  
 R;Kravartsi, L.; Lecanidou, R.; Rodakis, G.C.  
 submitted to the EMBL Data Library, February 1994  
 A;Description: Sequence analysis of a small early chorion gene subfamily interspersed  
 A;Reference number: S42070  
 A;Accession: S42070  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-168 <KEA>  
 A;Cross-references: EMBL:X77747; MID:G456246; PIDN:CAA54795.1; PID:G456247  
 C;Genetics:  
 A;Introns: 17/3  
 C;Superfamily: chorion class A protein pc292

Query Match 1.8%; Score 6; DB 2; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKIILF 161  
 |||||  
 Db 3 AKIILF 8

RESULT 210  
 B96916  
 hypothetical protein CAC0133 [imported] - Clostridium acetobutylicum  
 C;Species: Clostridium acetobutylicum  
 C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C;Accession: B96916  
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee  
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C  
 A;Reference number: A96900; MUID:21359325; PMID:21359325  
 A;Accession: B96916  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-168 <KUR>  
 A;Cross-references: GB:AE001437; PIDN:AAK78117.1; PID:G15022959; GSPDB:GN00168  
 A;Experimental source: Clostridium acetobutylicum ATCC824  
 C;Genetics:  
 A;Gene: CAC0133

Query Match 1.8%; Score 6; DB 2; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
 |||||  
 Db 7 VKILKD 12

RESULT 211  
 S23060  
 early chorion protein - silkworm  
 C;Species: Bombyx mori (silkworm)  
 C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C;Accession: S23060  
 R;Lecanidou, R.; Rodakis, G.C.  
 J. Mol. Evol. 34, 304-314, 1992  
 A;Title: Three copies of the early gene 6F6 are interspersed in and around the late ch  
 A;Reference number: S23060; MUID:92235869; PMID:1569584

A:Accession: S23060  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-169 <LEC>  
 A:Cross-references: EMBL:X66164; NID:g5737; PIDN:CAA46946.1; PID:g5738  
 C:Genetics:  
 A:Introns: 17/3  
 C:Superfamily: chorion class A protein pc292

Query Match 1.8%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKIILF 161  
 |||||  
 Db 3 AKIILF 8

RESULT 212  
 S42071  
 early chorion protein - silkworm  
 C:Species: Bombyx mori (silkworm)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Aug-1999  
 C:Accession: S42071  
 R:Kravartsi, L.; Lecanidou, R.; Rodakis, G.C.  
 submitted to the EMBL Data Library, February 1994  
 A:Description: Sequence analysis of a small early chorion gene subfamily interspersed with  
 A:Reference number: S42070  
 A:Accession: S42071  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-169 <KRA>  
 A:Cross-references: EMBL:X77746; NID:g456241; PIDN:CAA54794.1; PID:g456242  
 C:Genetics:  
 A:Introns: 17/3  
 C:Superfamily: chorion class A protein pc292

Query Match 1.8%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKIILF 161  
 |||||  
 Db 3 AKIILF 8

RESULT 213  
 C90175  
 NADH dehydrogenase subunit C (NuoC) NuoC [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C:Accession: C90175  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-  
 Jung, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: C90175  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-169 <KUR>  
 A:Cross-references: GB:AE006641; NID:g13813468; PIDN:AAK40658.1; GSPDB:GN00155  
 C:Genetics:  
 A:Gene: NuoC

Query Match 1.8%; Score 6; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160  
 |||||  
 Db 86 LAKIIL 91

RESULT 214  
 B75548  
 transcription regulator, Lrp/AsnC family - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: B75548  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; N  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: B75548  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-170 <WHI>  
 A:Cross-references: GB:AE001882; GB:AE000513; NID:g6457865; PIDN:AAF09787.1; PID:g64578  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0200  
 A:Map position: 1  
 C:Superfamily: regulatory protein asnC

Query Match 1.8%; Score 6; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELI 236  
 |||||  
 Db 138 LLGELI 143

RESULT 215  
 D81278  
 probable periplasmic protein Cjl343c [imported] - Campylobacter jejuni (strain NCTC 111  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: D81278  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre  
 Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: D81278  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-171 <PAR>  
 A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73770.1; PID:g69687  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cjl343c

Query Match 1.8%; Score 6; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVSKSL 46  
 |||||  
 Db 36 EVSKSL 41

RESULT 216  
 S27022  
 fibroblast growth factor receptor - Japanese medaka (fragment)  
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
 C:Species: Oryzias latipes (Japanese medaka)  
 C:Date: 19-Mar-1997 #sequence\_revision 24-Jul-1997 #text\_change 20-Jun-2000  
 C:Accession: S27022  
 R:Emori, Y.; Yasuoka, A.; Saigo, K.  
 FEBS Lett. 314, 176-178, 1992  
 A:Title: Identification of four FGF receptor genes in Medaka fish (Oryzias latipes).

A;Reference number: S27019; MUID:93093167; PMID:1459248  
 A;Accession: S27022  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-172 <EXO>  
 A;Cross-references: GB:DL1353; NID:G222938; PIDN:BAA02752.1; PID:G222939  
 C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote  
 F;1-172/Domain: alternative splicing; ATP; autophosphorylation; growth factor receptor; magn  
 F;23.40/132/Active site: Lys, Glu, Asp #status predicted  
 F;137.150/Binding site: magnesium (Asn, Asp) #status predicted  
 F;163/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 1.8%; Score 6; DB 2; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 186 TFXDLL 191  
 105 TFXDLL 110

RESULT 217  
 G64555  
 hypothetical protein HP0287 - Helicobacter pylori (strain 26695)  
 C;Species: Helicobacter pylori  
 C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 A;Accession: G64555  
 R;Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
 Nature 388, 539-547, 1997  
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A;Reference number: A6520; MUID:97394467; PMID:9252185  
 A;Accession: G64555  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-172 <TOM>  
 A;Cross-references: GB:AE000547; GB:AE000511; NID:G2313377; PIDN:RAD07359.1; PID:G231338

Query Match 1.8%; Score 6; DB 2; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 305 LSSFOK 310  
 74 LSSFOK 79

RESULT 218  
 F81122  
 hypothetical protein NMB1084 [imported] - Neisseria meningitidis (strain MC58 serogroup  
 C;Species: Neisseria meningitidis  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C;Accession: F81122  
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 Xia, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A;Reference number: A81000; MUID:20175755; PMID:10710307  
 A;Accession: F81122  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-173 <TET>  
 A;Cross-references: GB:AE002459; GB:AE002098; NID:G7226320; PIDN:AAF41476.1; PID:G722632  
 A;Superfamily: serogroup B, strain MC58  
 C;Genetics:  
 A;Gene: NMB1084  
 C;Superfamily: Neisseria meningitidis hypothetical protein NMB1084

Query Match 1.8%; Score 6; DB 2; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 317 QFADEK 322  
 51 QFADEK 56

RESULT 219  
 F70112  
 hypothetical protein BB0102 - Lyme disease spirochete  
 C;Species: Borrelia burgdorferi (lyme disease spirochete)  
 C;Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999  
 C;Accession: F70112  
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit  
 son, D.; Peterson, J.; Kervavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt  
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A;Authors: Smith, H.O.; Venter, J.C.  
 A;Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
 A;Reference number: A70100; MUID:98065943; PMID:9403685  
 A;Accession: F70112  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-173 <KLE>  
 A;Cross-references: GB:AB001123; GB:AE000783; NID:G2687987; PIDN:AAC66505.1; PID:G26880  
 A;Experimental source: strain B31

Query Match 1.8%; Score 6; DB 2; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 20 ILKDNL 25  
 146 ILKDNL 151

RESULT 220  
 A35383  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Caulobacter crescentus  
 C;Species: Caulobacter crescentus  
 C;Date: 31-Aug-1990 #sequence\_revision 31-Aug-1990 #text\_change 18-Jul-2001  
 C;Accession: A35383; B87445  
 R;Steinman, H.M.; Ely, B.  
 J. Bacteriol. 172, 2901-2910, 1990  
 A;Title: Copper-zinc superoxide dismutase of Caulobacter crescentus: cloning, sequencin  
 A;Reference number: A35383; MUID:90264275; PMID:2345128  
 A;Accession: A35383  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-174 <STE>  
 A;Cross-references: GB:M55259; NID:G144282; PIDN:AAA23054.1; PID:G144283; GB:M37908  
 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A;Title: Complete Genome Sequence of Caulobacter crescentus.  
 A;Reference number: A87249; MUID:21173698; PMID:11259647  
 A;Accession: B87445  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-174 <STO>  
 A;Cross-references: GB:AE005673; NID:G13422970; PIDN:AAK23558.1; GSPDB:GN00148  
 C;Genetics:  
 A;Gene: CC1579  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase (Cu-Zn)  
 C;Keywords: metalloprotein; oxidoreductase  
 F;167/Active site: Arg #status predicted

Query Match 1.8%; Score 6; DB 2; Length 174;



Best Local Similarity 100.0%; Pred. No. 3.8e+02; Mismatches 0; Gaps 0;  
 Matches 6; Conservative 0; Indels 0;

QY 285 HKTQPI 290  
 |||||  
 Db 156 HKTQPI 161

RESULT 221  
 F75606  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: F75606  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: F75606  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-174 <NHI>  
 A:Cross-references: GB:AE001862; GB:AE001825; NID:G6460468; PIDN:AAF12321.1; PID:G646061  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DRA0116  
 A:Map position: 2

Query Match 1.8%; Score 6; DB 2; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 78 LIAVTLI 83  
 |||||  
 Db 81 LIAVTLI 86

RESULT 222  
 E90006  
 hypothetical protein rpoE [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 27-Oct-2003  
 C:Accession: E90006  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89759; MUID:21311952; PMID:11418146  
 A:Accession: E90006  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-176 <KUR>  
 A:Cross-references: GB:BA000018; PID:G13701922; PIDN:BA043214.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: rpoE  
 C:Superfamily: DNA-directed RNA polymerase delta subunit

Query Match 1.8%; Score 6; DB 2; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 229 LKILGE 234  
 |||||  
 Db 111 LKILGE 116

RESULT 223  
 T02217  
 NBS-LRR type resistance protein - rice (fragment)

C:Species: Oryza sativa (rice)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 21-Jul-2000  
 C:Accession: T02217  
 R:Leister, D.; Kurth, J.; Laurie, D.A.; Yano, M.; Sasaki, T.; Devos, K.; Graner, A.; Scl  
 Proc. Natl. Acad. Sci. U.S.A. 95, 370-375, 1998  
 A:Title: Rapid reorganization of resistance gene homologues in cereal genomes.  
 A:Reference number: Z14623; MUID:98081880; PMID:9419382  
 A:Accession: T02217  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-176 <LEI>  
 A:Cross-references: EMBL:AF032691; NID:G2792225; PIDN:BA096988.1; PID:G2792226  
 A:Experimental source: subsp. Japonica, cultivar. Nipponbare  
 C:Genetics:  
 A:Gene: r4

Query Match 1.8%; Score 6; DB 2; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 292 EILLKN 297  
 |||||  
 Db 65 EILLKN 70

RESULT 224  
 AB0414  
 hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) [imported] - Yersinia pestis (strain  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AB0414  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 II, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AB0414  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-178 <KUR>  
 A:Cross-references: GB:AL50842; PIDN:CAC92638.1; PID:G15981335; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: hpt  
 C:Superfamily: hypoxanthine phosphoribosyltransferase  
 C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 1.8%; Score 6; DB 2; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;

QY 18 VKILKD 23  
 |||||  
 Db 80 VKILKD 85

RESULT 225  
 AH0523  
 hypoxanthine phosphoribosyltransferase [imported] - Salmonella enterica subsp. enterica  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AH0523  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher  
 Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AH0523  
 A:Status: preliminary  
 A:Molecule type: DNA

A;Residues: 1-178 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD01328.1; PID:g16501456; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STY0192  
 C;Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 1.8%; Score 6; DB 2; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
 |||||  
 DB 80 VKILKD 85

## RESULT 226

T23601

hypothetical protein K11D2.1 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T23601

R;Kershaw, J.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19768

A;Accession: T23601

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-178 &lt;WIL&gt;

A;Cross-references: EMBL:Z83115; PIDN:CAB05555.1; GSPDB:GN00019; CBSP:K11D2.1  
 A;Experimental source: clone K11D2

C;Genetics:

A;Gene: CBSP:K11D2.1

A;Map position: 1

A;Introns: 109/2

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 178;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKIL 21

|||||

DB 53 EIVKIL 58

## RESULT 227

E70358

HupE hydrogenase related function - Aquifex aeolicus  
 C;Species: Aquifex aeolicus

A;Reference number: A70300; MUID:98196666; PMID:9537320

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 28-Jul-2003  
 C;Accession: E70358

R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: E70358

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-180 &lt;AQF&gt;

A;Cross-references: GB:AE000701; NID:G2983260; PIDN:AAC06859.1; PID:G2983265; GB:AF00065

A;Experimental source: strain VFS

C;Genetics:

A;Gene: hupE

C;Superfamily: [NiFe]-hydrogenase/urease accessory protein HupE/UreJ

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 180;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GYEAPQ 137

|||||

DB 127 GYEAPQ 132

## RESULT 228

AI0861

syd protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT16  
 C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AI0861

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AI0861

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-181 &lt;PAR&gt;

A;Cross-references: GB:AL513382; PIDN:CAD06080.1; PID:g16504046; GSPDB:GN00176

C;Genetics:

A;Gene: STY3106

C;Superfamily: Escherichia coli syd protein

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 181;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 QFADEK 322

|||||

DB 92 QFADEK 97

## RESULT 229

T07837

thioredoxin f precursor - rape  
 C;Species: Brassica napus (rape)

C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jan-2000  
 C;Accession: T07837

R;Mora-Garcia, S.E.F.; Rodriguez-Suarez, R.J.; Wolosiuk, R.A.

submitted to the EMBL Data Library, August 1997

A;Reference number: Z16164

A;Accession: T07837

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-182 &lt;MOR&gt;

A;Cross-references: EMBL:AF018174; NID:G2921093; PIDN:AAC04671.1; PID:G2921094

A;Experimental source: 30-day old green leaves

C;Genetics:

A;Gene: TRXF

A;Genome: nuclear

C;Function:

A;Description: activates fructose 1,6-bisphosphatase during photosynthesis

C;Superfamily: thioredoxin; thioredoxin homology

C;Keywords: chloroplast; redox-active disulfide

F;779-163/Domain: thioredoxin homology &lt;TXN&gt;

## Query Match

Best Local Similarity 1.8%; Score 6; DB 2; Length 182;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILKDN 24

|||||

DB 149 KILKDN 154

## RESULT 230

S20929

thioredoxin f precursor - garden pea  
 C;Species: Pisum sativum (garden pea)

C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 11-Jun-1999  
 C;Accession: S20929

R;Lepintec, L.; Hodges, M.; Gadgil, P.; Cretin, C.

Plant Mol. Biol. 18, 1023-1025, 1992  
 A:Title: Isolation, characterization and nucleotide sequence of a full-length pea cDNA  
 A:Reference number: S20929; MUID:92256804; PMID:1581563  
 A:Accession: S20929  
 A:Molecule type: mRNA  
 A:Residues: 1-182 <LEP>  
 A:Cross-references: EMBL:X63537; NID:g290906; PIDN:CAR45098.1; PID:g20907  
 C:Superfamily: thioredoxin; thioredoxin homology  
 C:Keywords: chloroplast; redox-active disulfide  
 F:1-74/Domain: transit peptide (chloroplast) #status predicted <TNP>  
 F:75-182/Product: thioredoxin f #status predicted <MAT>  
 F:83-167/Domain: thioredoxin homology <HR>  
 F:106-109/Disulfide bonds: redox-active #status predicted

Query Match 1.8%; Score 6; DB 2; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILKDN 24  
 Db 153 KILKDN 158  
 |||||

RESULT 231  
 A90645  
 hypoxanthine phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, su  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C:Accession: A90645  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
 A:Reference number: A90629; MUID:21156231; PMID:11258796  
 A:Accession: A90645  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-182 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA33552.1; PID:gl3359585; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: ECs0129  
 C:Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 1.8%; Score 6; DB 2; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
 Db 84 VKILKD 89  
 |||||

RESULT 232  
 A85496  
 hypoxanthine phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, su  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: A85496  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: A85496  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-182 <STO>  
 A:Cross-references: GB:AB005174; NID:gl2512840; PIDN:AGS4429.1; GSPDB:GN00145; UWGP:Z01  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: hpt  
 C:Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 1.8%; Score 6; DB 2; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
 Db 84 VKILKD 89  
 |||||

RESULT 233  
 hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 01-Mar-2002  
 C:Accession: S45202; E64735  
 R:Fujita, N. The EMBL Data Library, January 1994  
 A:Reference number: S45181  
 A:Accession: S45202  
 A:Molecule type: DNA  
 A:Residues: 1-182 <FUJ>  
 A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BAA05581.1; PID:g473792  
 A:Experimental source: strain K-12, substrain W3110  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: E64735  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-182 <BLAT>  
 A:Cross-references: GB:AF000122; GB:U00096; NID:gl786315; PIDN:AACT3236.1; PID:gl786317  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: hpt  
 C:Superfamily: hypoxanthine phosphoribosyltransferase  
 C:Keywords: glycosyltransferase; pentosyltransferase; salvage pathway

Query Match 1.8%; Score 6; DB 2; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
 Db 84 VKILKD 89  
 |||||

RESULT 234  
 F71072  
 hypothetical protein PH1273 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: F71072  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: F71072  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-182 <KAW>  
 A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BAA30376.1; PID:g3257693  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH1273

Query Match 1.8%; Score 6; DB 2; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 VTLIAD 85  
DB 125 VTLIAD 130

RESULT 235  
F70509  
Hypothetical protein Rv0513 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: F70509  
R:Colts, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: F70509  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-182 <COL>  
A:Cross-references: GB:Z97831; GB:AL123456; NID:G3261825; PIDN:CAB10750.1; PID:e329664;  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: Rv0513

Query Match 1.8%; Score 6; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKI 20  
DB 90 AEIVKI 95

RESULT 236  
G71921  
Hypothetical protein jhp0526 - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
C:Accession: G71921  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Moberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: G71921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-183 <ARN>  
A:Cross-references: GB:AE001485; GB:AE001439; NID:G4155057; PIDN:AA06107.1; PID:G415507  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: jhp0526

Query Match 1.8%; Score 6; DB 2; Length 183;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 EFLSSF 308  
DB 52 EFLSSF 57

RESULT 237  
C85439  
probable cytoskeletal protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C85439

Query Match 1.8%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIE 303  
DB 148 QPKLIE 153

RESULT 238  
C64592  
Hypothetical protein HP0579 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: C64592  
R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: C64592  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-184 <TOM>  
A:Cross-references: GB:AE000571; GB:AE000511; NID:G2313686; PIDN:AA07645.1; PID:G23136

Query Match 1.8%; Score 6; DB 2; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 EFLSSF 308  
DB 52 EFLSSF 57

RESULT 239  
S27936  
Hypothetical protein 1 - sugarcane bacilliform virus  
C:Species: sugarcane bacilliform virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S27936  
R:Bouhida, M.; Lockhart, B.; Olszewski, N.E.  
Submitted to the EMBL Data Library, May 1992  
A:Description: The complete sequence of an infective sugarcane mosaic virus genome.  
A:Reference number: S27936  
A:Accession: S27936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <BOU>  
A:Cross-references: EMBL:M89923; NID:G294836; PIDN:AAA47452.1; PID:G294837

Query Match 1.8%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIE 303  
DB 148 QPKLIE 153

RESULT 239  
S27936  
Hypothetical protein 1 - sugarcane bacilliform virus  
C:Species: sugarcane bacilliform virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S27936  
R:Bouhida, M.; Lockhart, B.; Olszewski, N.E.  
Submitted to the EMBL Data Library, May 1992  
A:Description: The complete sequence of an infective sugarcane mosaic virus genome.  
A:Reference number: S27936  
A:Accession: S27936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <BOU>  
A:Cross-references: EMBL:M89923; NID:G294836; PIDN:AAA47452.1; PID:G294837

Query Match 1.8%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIE 303  
DB 148 QPKLIE 153

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: C85439  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-183 <STO>  
A:Cross-references: GB:NC\_001268; NID:G2720667; PIDN:CAB80384.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4G37180  
A:Map position: 4

Query Match 1.8%; Score 6; DB 2; Length 183;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 KOIRDL 332  
DB 66 KOIRDL 71

RESULT 238  
C64592  
Hypothetical protein HP0579 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: C64592  
R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: C64592  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-184 <TOM>  
A:Cross-references: GB:AE000571; GB:AE000511; NID:G2313686; PIDN:AA07645.1; PID:G23136

Query Match 1.8%; Score 6; DB 2; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 EFLSSF 308  
DB 52 EFLSSF 57

RESULT 239  
S27936  
Hypothetical protein 1 - sugarcane bacilliform virus  
C:Species: sugarcane bacilliform virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S27936  
R:Bouhida, M.; Lockhart, B.; Olszewski, N.E.  
Submitted to the EMBL Data Library, May 1992  
A:Description: The complete sequence of an infective sugarcane mosaic virus genome.  
A:Reference number: S27936  
A:Accession: S27936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <BOU>  
A:Cross-references: EMBL:M89923; NID:G294836; PIDN:AAA47452.1; PID:G294837

Query Match 1.8%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIE 303  
DB 148 QPKLIE 153

RESULT 238  
C64592  
Hypothetical protein HP0579 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: C64592  
R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: C64592  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-184 <TOM>  
A:Cross-references: GB:AE000571; GB:AE000511; NID:G2313686; PIDN:AA07645.1; PID:G23136

Query Match 1.8%; Score 6; DB 2; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 EFLSSF 308  
DB 52 EFLSSF 57

RESULT 239  
S27936  
Hypothetical protein 1 - sugarcane bacilliform virus  
C:Species: sugarcane bacilliform virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S27936  
R:Bouhida, M.; Lockhart, B.; Olszewski, N.E.  
Submitted to the EMBL Data Library, May 1992  
A:Description: The complete sequence of an infective sugarcane mosaic virus genome.  
A:Reference number: S27936  
A:Accession: S27936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <BOU>  
A:Cross-references: EMBL:M89923; NID:G294836; PIDN:AAA47452.1; PID:G294837

Query Match 1.8%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIE 303  
DB 148 QPKLIE 153

RESULT 239  
S27936  
Hypothetical protein 1 - sugarcane bacilliform virus  
C:Species: sugarcane bacilliform virus  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S27936  
R:Bouhida, M.; Lockhart, B.; Olszewski, N.E.  
Submitted to the EMBL Data Library, May 1992  
A:Description: The complete sequence of an infective sugarcane mosaic virus genome.  
A:Reference number: S27936  
A:Accession: S27936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <BOU>  
A:Cross-references: EMBL:M89923; NID:G294836; PIDN:AAA47452.1; PID:G294837

## RESULT 240

envelope-associated 22K protein - turkey rhinotracheitis virus  
 C:Species: turkey rhinotracheitis virus  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C:Accession: JQ1623  
 R: Ling, R.; Easton, A.J.; Pringle, C.R.  
 J. Gen. Virol. 73, 1709-1715, 1992  
 A: Title: Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus and  
 A: Reference number: PQ0405; MUID: 92333255; PMID: 1629697  
 A: Accession: JQ1623  
 A: Molecule type: mRNA  
 A: Residues: 1-186 <L>IN>  
 A: Cross-references: GB:S40185; NID:G251600; PIDN: AAB22544.1; PID: G251602  
 C: Genes: 22K  
 C: Superfamily: respiratory syncytial virus envelope-associated 22K protein  
 C: Keywords: glycoprotein  
 F: 86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 6; DB 1; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160  
 |||||  
 Db 158 LAKIIL 163

## RESULT 241

hypothetical 21.0K protein - turkey rhinotracheitis virus (strain UK/3BV/85)  
 N: Alternate names: ORF 1 protein  
 C: Species: turkey rhinotracheitis virus  
 C: Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 28-May-1999  
 C: Accession: JQ1987  
 R: Yu, Q.; Davis, P.J.; Brown, T.D.K.; Cavanagh, D.  
 J. Gen. Virol. 73, 1355-1363, 1992  
 A: Title: Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pneum  
 A: Reference number: JQ1987; MUID: 92300329; PMID: 1607858  
 A: Accession: JQ1987  
 A: Molecule type: mRNA  
 A: Residues: 1-186 <YUQ>  
 A: Cross-references: GB:X63408; NID: G297846; PIDN: CAA45004.1; PID: G297847  
 C: Superfamily: respiratory syncytial virus envelope-associated 22K protein

Query Match 1.8%; Score 6; DB 2; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160  
 |||||  
 Db 158 LAKIIL 163

## RESULT 242

late competence operon required for DNA binding and uptake comEB [imported] - Bacillus h  
 F83816  
 C: Species: Bacillus halodurans  
 C: Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C: Accession: F83816  
 R: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A: Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A: Reference number: A83650; MUID: 20512582; PMID: 11058132  
 A: Accession: F83816  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-188 <STO>  
 A: Cross-references: GB:AP001511; GB:BA000004; NID: g10173727; PIDN: BAB05053.1; GSPDB: GN00

Query Match 1.8%; Score 6; DB 2; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ISKPEN 254  
 |||||  
 Db 142 ISKPEN 147

## C:Genetics:

A: Gene: comEB

C: Superfamily: phage T4 dCMP deaminase

Query Match 1.8%; Score 6; DB 2; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 ILDRHN 241  
 |||||  
 Db 145 ILDRHN 150

## RESULT 243

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) [imported] - Brucella melitensis (strain  
 C: Species: Brucella melitensis  
 C: Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-May-2002  
 C: Accession: AG3395  
 R: DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzeck, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A: Title: The genome sequence of the facultative intracellular pathogen Brucella meliten  
 A: Reference number: AD3252; PMID: 11756688  
 A: Accession: AG3395  
 A: Status: preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-189 <KUR>  
 A: Cross-references: GB:AE008917; PIDN: AAL52330.1; PID: g17983124; GSPDB: GN00190  
 A: Experimental source: strain 16M  
 C: Genetics:

A: Gene: BME11149  
 A: Map position: 1  
 C: Superfamily: NADH dehydrogenase (ubiquinone) chain 6  
 C: Keywords: oxidoreductase

Query Match 1.8%; Score 6; DB 2; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELI 236  
 |||||  
 Db 88 LLGELI 93

## RESULT 244

G89581

protein C18A11.4 [imported] - Caenorhabditis elegans

C: Species: Caenorhabditis elegans

C: Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001

C: Accession: G89581

R: anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998  
 A: Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
 A: Reference number: A75000; MUID: 99069613; PMID: 9851916

A: Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/c\_el  
 A: Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an

A: Accession: G89581

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-190 <STO>

A: Cross-references: GB:chr\_X; PIDN: AAC69008.1; PID: g1049416; GSPDB: GN000028; CESP: C18A11

C: Genetics:

A: Gene: C18A11.4

A: Map position: X

Query Match 1.8%; Score 6; DB 2; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ISKPEN 254  
 |||||  
 Db 142 ISKPEN 147

RESULT 245  
G70940  
hypothetical protein rv2016 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: G70940  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70940  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-191 <COL>  
A:Cross-references: GB:AL021899; GB:AL123456; NID:G3242282; PIDN:CAAL7230.1; PID:G289679  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV2016  
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2016

Query Match 1.8%; Score 6; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 LVTLIA 84  
|||||  
Db 54 LVTLIA 59

RESULT 246  
C81279  
probable molybdopterin-guanine dinucleotide biosynthesis protein Cj1350 [imported] - Cam  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: C81279  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: AB1250; MUID:20150912; PMID:10686204  
A:Accession: C81279  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-191 <PAR>  
A:Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73777.1; PID:G696878  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: mobA; Cj1350

Query Match 1.8%; Score 6; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 FLEQNY 206  
|||||  
Db 112 FLEQNY 117

RESULT 247  
S76506  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S76506  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S76506  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-192 <KAN>  
A:Cross-references: EMBL:D64002; GB:AB001339; NID:G1001612; PIDN:BAAL0352.1; PID:G10016;  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: conserved hypothetical protein slr2127

Query Match 1.8%; Score 6; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LLGELI 236  
|||||  
Db 11 LLGELI 16

RESULT 248  
AD2488  
hypothetical protein all7084 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AD2488  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi;  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD2488  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-192 <KUR>  
A:Cross-references: GB:BA000020; PIDN:BA078168.1; PID:G17135622; GSPDB:GN00180  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all7084  
A:Genome: plasmid

Query Match 1.8%; Score 6; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EIVKIL 21  
|||||  
Db 43 EIVKIL 48

RESULT 249  
AB4191  
3-dehydroquinate dehydratase [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 07-Jul-2003  
C:Accession: AB4191  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1  
A:Reference number: AB4160; MUID:20504483; PMID:11016950  
A:Accession: AB4191  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-195 <STO>  
A:Cross-references: GB:AB004437; NID:G10579943; PIDN:AAG18893.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: arod  
C:Superfamily: 3-dehydroquinate dehydratase; 3-dehydroquinate dehydratase homology

Query Match 1.8%; Score 6; DB 2; Length 195;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LIADLQ 87  
Db 190 LIADLQ 195

RESULT 250  
JC7106  
development related unidentified 25K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000  
C:Accession: JC7106  
R:Sakuma-Takagi, M.; Tohyama, Y.; Kasama-Yoshida, H.; Sakagami, H.; Kondo, H.; Kurihara,  
Biochem. Biophys. Res. Commun. 263, 737-742, 1999  
A:Title: Novel related cDNAs (C184L, C184M, and C184S) from developing mouse brain encod  
A:Reference number: JC7106; MUID:99443733; PMID:10512749  
A:Accession: JC7106  
A:Molecule type: mRNA  
A:Residues: 1-199 <SAK>  
A:Cross-references: DDBJ:AB012727  
A:Experimental source: brain, C184S  
C:Keywords: brain; zinc finger

Query Match 1.8%; Score 6; DB 2; Length 199;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 EPTTEA 65  
Db 14 EPTTEA 19

Search completed: April 12, 2004, 10:37:19  
Job time : 30 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:29:13 ; Search time 18 Seconds  
(without alignments)

974.869 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 337

Sequence: 1 MKMPLFSKSHKNPABIVK1.....FADEKNYLIKQIRDLKKTAP 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	233	69.1	334	1	Q919s4 homo sapien
2	121	35.9	334	1	Q9db16 mus musculus
3	23	6.8	341	1	Q9y376 homo sapien
4	23	6.8	341	1	Q06138 mus musculus
5	16	4.7	321	1	Q9xfy6 chlorella p
6	16	4.7	329	1	Q9p7q8 schizosacch
7	15	4.5	339	1	P91891 drosophila
8	14	4.2	338	1	O18211 caenorhabdi
9	9	2.7	384	1	O60032 emericella
10	8	2.4	923	1	P18475 drosophila
11	7	2.1	139	1	Q8xrm1 raltstonia s
12	7	2.1	156	1	Q60313 methanococc
13	7	2.1	177	1	O51875 buchnera ap
14	7	2.1	228	1	P26697 gallus gall
15	7	2.1	232	1	Q8xf15 fusobacteri
16	7	2.1	293	1	P77044 escherichia
17	7	2.1	339	1	P21358 candida gla
18	7	2.1	340	1	P17576 polyporus t
19	7	2.1	343	1	Q9m0m4 arabidopsis
20	7	2.1	343	1	Q8fgk3 arabidopsis
21	7	2.1	358	1	Q8fve4 brucella su
22	7	2.1	359	1	Q59175 brucella ab
23	7	2.1	359	1	Q8ycy1 brucella me
24	7	2.1	372	1	P47459 mycoplasma
25	7	2.1	373	1	O66875 aquifex aeo
26	7	2.1	399	1	P32464 saccharomyc
27	7	2.1	421	1	Q02728 rhizobium m
28	7	2.1	423	1	P53663 mycoplasma
29	7	2.1	430	1	Q9sfh9 arabidopsis
30	7	2.1	438	1	GLNA RHOC
31	7	2.1	474	1	YB8G YEAST
32	7	2.1	490	1	MYOC MOUSE
33	7	2.1	492	1	P79739 brachydanio

34	7	2.1	500	1	RASP DROME
35	7	2.1	502	1	MYOC RAT
36	7	2.1	504	1	MYOC HUMAN
37	7	2.1	535	1	C4C3_DROME
38	7	2.1	762	1	PMIP_SCHPO
39	7	2.1	923	1	PH87 YEAST
40	7	2.1	1237	1	NME3 RAT
41	7	2.1	1239	1	NME3 MOUSE
42	7	2.1	1489	1	RNG2_SCHPO
43	7	2.1	2869	1	RBPI_PLAVB
44	7	2.1	3433	1	UTRO HUMAN
45	7	2.1	3788	1	LYST MOUSE
46	6	1.8	15	1	FIBA_ANAPL
47	6	1.8	38	1	BD08_BOVIN
48	6	1.8	40	1	BD07_BOVIN
49	6	1.8	55	1	BD09_BOVIN
50	6	1.8	72	1	RL29_TREPA
51	6	1.8	82	1	Y056_NPOVP
52	6	1.8	83	1	Y46S_SVNY3
53	6	1.8	87	1	RL31_METJA
54	6	1.8	90	1	41_CHICK
55	6	1.8	101	1	4BP3_MOUSE
56	6	1.8	101	1	VNS7_CVCAE
57	6	1.8	110	1	Y18K_MSVK
58	6	1.8	110	1	YCX1_CHLPY
59	6	1.8	113	1	SENA_APLCA
60	6	1.8	124	1	Y905_ARCFU
61	6	1.8	128	1	YC35_CYAPA
62	6	1.8	131	1	RS6E_HALN1
63	6	1.8	133	1	RS8_CHLPN
64	6	1.8	134	1	RL27_PVRST
65	6	1.8	134	1	RR9_EUGGR
66	6	1.8	135	1	IF2B_METH
67	6	1.8	135	1	YJAV_BACSU
68	6	1.8	138	1	XR8_OENHO
69	6	1.8	140	1	IF2B_PFRAB
70	6	1.8	140	1	IF2B_PYRHO
71	6	1.8	140	1	IF2B_PYRHO
72	6	1.8	142	1	U426_HSVMG
73	6	1.8	143	1	HS16_CAEEL
74	6	1.8	145	1	RL9_UREFA
75	6	1.8	150	1	TAA1_VACCC
76	6	1.8	150	1	TAA1_VACCV
77	6	1.8	150	1	TAA1_VAV
78	6	1.8	151	1	RVUX_ANASP
79	6	1.8	152	1	SSRP_HELPJ
80	6	1.8	152	1	SSRP_HELPJ
81	6	1.8	153	1	Y18K_MSVN
82	6	1.8	155	1	IL2_BOVIN
83	6	1.8	155	1	IL2_BUBBU
84	6	1.8	155	1	IL2_SHEEP
85	6	1.8	161	1	RL10_MYCPE
86	6	1.8	161	1	RL10_MYCPN
87	6	1.8	162	1	IL2_CEREL
88	6	1.8	162	1	RL10_MYCCE
89	6	1.8	162	1	YP92_CAEEL
90	6	1.8	165	1	RS16_CORGL
91	6	1.8	170	1	GVA2_STRCO
92	6	1.8	174	1	SODC_CAUCR
93	6	1.8	176	1	CYB_NYCHU
94	6	1.8	176	1	RPOE_STAMM
95	6	1.8	178	1	HPRT_ECOLI
96	6	1.8	178	1	HPRT_SALTY
97	6	1.8	178	1	THF1_ARATH
98	6	1.8	179	1	Y969_VIBPA
99	6	1.8	181	1	SYDP_SALTY
100	6	1.8	181	1	YHCE_ECOLI
101	6	1.8	181	1	THIF_BRANA
102	6	1.8	182	1	THIF_PEA
103	6	1.8	182	1	EPF_BACFR
104	6	1.8	185	1	THF2_ARATH
105	6	1.8	185	1	VMA2_TRTV
106	6	1.8	186	1	

Q9vzu2 drosophila	Q9rlj4 rattus norv	Q99572 homo sapien	Q9va27 drosophila	Q10415 schizosacch	P25360 saccharomyc	Q00961 rattus norv	Q01098 mus musculu	Q14188 schizosacch	Q00798 plasmodium	P46939 homo sapien	P97412 mus musculu	P12801 anas platyr	P46166 bos taurus	P46165 bos taurus	P46167 bos taurus	Q83227 treponema p	O10314 orgyia pseu	P73882 synchocyst	P54009 methanococc	P12264 gallus gall	Q80v3 mus musculu	Q04703 canine ente	P14988 maize strea	P05720 chlorella p	P29233 aplysia cal	O29357 archaeoglob	P48275 cyanophora	Q9hmj5 halobacteri	Q92794 chlamydia p	Q02984 pyrobactrys	P32060 euglena gra	O27797 methanobact	O34959 bacillus su	Q9mctj0 oenochera h	Q9uyr6 pyrococcus	Q8u315 pyrococcus	O58312 pyrococcus	O51014 marek's dis	P06581 caenorhabdi	Q9p8t9 ureaplasma	Q20982 vaccinia vi	P07610 vaccinia vi	P33814 variola vir	O8yis5 anabaena sp	Q8zjh2 helicobacte	O25985 helicobacte	P14978 maize strea	P05016 bos taurus	Q95kp3 bubalus bub	P19114 ovibis aries	O8evj0 mycoplasma	P5240 mycoplasma	P51747 cervus elap	Q36263 mycoplasma	Q09243 caenorhabdi	Q8nnx3 corynebacte	Q9rjb4 streptomyce	P20379 caulobacter	Q36572 nysticeus	Q99sd0 staphylococ	P36766 escherichia	O33799 salmoneilla	Q9xfh8 arabidopsis	Q8r729 vibrio para	P60085 salmoneilla	P60084 salmoneilla	P45421 escherichia	O48897 brassica na	P29450 pisum sativ	P70889 bacteroides	Q9xfh9 arabidopsis	P33494 turkey rhin
-------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	-------------------	-------------------	-------------------	--------------------	--------------------	-------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	---------------------	-------------------	-------------------	-------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	---------------------	-------------------	------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------

107	6	1.8	131	1	MOBA_CAMJE	Q9pmu9 campylobact	180	6	1.8	291	1	HSLO_VIBPA	Q87le0 vibrio para
108	6	1.8	138	1	Y4D2_METAC	Q8tiv9 methanosarc	181	6	1.8	291	1	HSLO_VIBPA	Q8ds7 vibrio vuln
109	6	1.8	139	1	AA27_HUMAN	O60232 homo sapien	182	6	1.8	292	1	PHFB_HUMAN	Q9u18 homo sapien
110	6	1.8	139	1	AA27_MOUSE	P56873 mus musculu	183	6	1.8	292	1	TRUB_STRPN	Q97q13 streptococ
111	6	1.8	200	1	HAMI_CAMJE	O9pm56 campylobact	184	6	1.8	292	1	TRUB_STRPN	Q8cur2 streptococ
112	6	1.8	200	1	HIS5_METKA	Q8tv93 methanopyru	185	6	1.8	295	1	SU01_RAT	P52844 rattus norv
113	6	1.8	201	1	YMW3_YEAST	Q4772 saccharomyc	186	6	1.8	295	1	SU02_RAT	P52845 rattus norv
114	6	1.8	202	1	TRPF_BACCR	Q81996 bacillus ce	187	6	1.8	295	1	SU03_RAT	P49889 rattus norv
115	6	1.8	203	1	TFAB_ECOLI	Q47427 escherichia	188	6	1.8	295	1	SU06_RAT	P49890 rattus norv
116	6	1.8	203	1	TRPF_YEAST	P40080 saccharomyc	189	6	1.8	295	1	SU01_MOUSE	P49891 mus musculu
117	6	1.8	204	1	TRPF_BACAA	O81t19 bacillus an	190	6	1.8	295	1	XERC_PASMU	Q9kx2 pasteurella
118	6	1.8	206	1	ICT1_HUMAN	O14197 homo sapien	191	6	1.8	296	1	ERA_FUSNN	Q8gm1 fusobacteri
119	6	1.8	206	1	PCP2_THETN	Q89j6 thermotanaer	192	6	1.8	296	1	TYST_AGABI	Q9pt7 agaricus bi
120	6	1.8	207	1	YIOR_CVBF	Z22654 bovine coro	193	6	1.8	296	1	HEM3_FUSNN	Q8rip5 fusobacteri
121	6	1.8	209	1	IF2B_THEAC	O9hkj3 thermoplasm	194	6	1.8	298	1	RT03_ACACA	P46754 acanthamoeb
122	6	1.8	210	1	GTP2_BUBBU	P83325 bufo bufo (	195	6	1.8	298	1	XERD_LEPIN	Q7zaw7 leptospira
123	6	1.8	210	1	HIS1_SYNY3	O55503 synecocyst	196	6	1.8	299	1	G3P_DICDI	Q94469 dictyosteli
124	6	1.8	211	1	CCGX_MOUSE	O9jiv3 mus musculu	197	6	1.8	299	1	HIS1_BUCBP	P59453 buchnera ap
125	6	1.8	211	1	MOTA_BPT4	P22915 bacterioph	198	6	1.8	299	1	HSLO_SYNEL	Q8dkq1 synecococ
126	6	1.8	212	1	PUR3_HAEIN	P43846 haemophilus	199	6	1.8	300	1	FTRL_METKA	Q8tX60 methanopyru
127	6	1.8	212	1	RADB_METJA	Q57702 methanococ	200	6	1.8	301	1	SC14_KLULA	P24859 kluyveromyc
128	6	1.8	217	1	MMCI_METAC	P58867 methanosarc	201	6	1.8	302	1	YF85_MYCPN	P75195 mycoplasma
129	6	1.8	217	1	YB63_METJA	O58563 methanococ	202	6	1.8	304	1	Y191_METTH	O26293 methanobact
130	6	1.8	218	1	FTSE_HAEIN	P44871 haemophilus	203	6	1.8	305	1	NUIM_APILI	P34847 apis mellif
131	6	1.8	220	1	ENGB_VIRCH	Q9kvn0 vibrio chol	204	6	1.8	305	1	PEEL_SCHPO	P36614 schizosacch
132	6	1.8	221	1	KTHV_LACPL	Q88yp6 lactobacill	205	6	1.8	307	1	TRUB_HELMO	Q86999 heliobacill
133	6	1.8	222	1	FTSE_ECOLI	P10115 escherichia	206	6	1.8	308	1	MRAW_HELPY	O25411 heliobacill
134	6	1.8	222	1	GL14_ARATH	O9m8x4 arabidopsis	207	6	1.8	309	1	DHL2_LACCO	P14295 lactobacill
135	6	1.8	222	1	GL15_ARATH	O9m8x5 arabidopsis	208	6	1.8	309	1	O4C6_HUMAN	Q8n7t3 homo sapien
136	6	1.8	223	1	RPL1_BPD3	Q37906 bacterioph	209	6	1.8	311	1	DHYS_SULTO	O971t3 sulfobolus
137	6	1.8	224	1	PLF1_MOUSE	P04095 mus musculu	210	6	1.8	311	1	NADA_SULSO	Q972c4 sulfobolus
138	6	1.8	224	1	PLF2_MOUSE	P04768 mus musculu	211	6	1.8	311	1	SRG6_CAEEL	P54128 caenorhabdi
139	6	1.8	224	1	PLF3_MOUSE	P18918 mus musculu	212	6	1.8	313	1	HEM3_PSESM	Q88D91 pseudomonas
140	6	1.8	226	1	AROD_HALN1	Q9hsb4 halobacteri	213	6	1.8	314	1	DHBV_CAEEL	Q17703 caenorhabdi
141	6	1.8	226	1	RADC_STAEP	Q8cnz4 staphylococ	214	6	1.8	314	1	O5BH_HUMAN	Q8n5f7 homo sapien
142	6	1.8	226	1	VATE_MANSE	P31402 manduca sex	215	6	1.8	314	1	PRMA_LISMO	O955a2 listeria mo
143	6	1.8	227	1	GL13_ARATH	O9m8x3 arabidopsis	216	6	1.8	315	1	DHBX_CAEEL	O17795 caenorhabdi
144	6	1.8	232	1	Y685_METTH	O26781 methanobact	217	6	1.8	315	1	MFTC_HUMAN	Q9hxd1 homo sapien
145	6	1.8	235	1	PSB3_DROME	P40304 drosophila	218	6	1.8	315	1	MFTC_MACFA	Q95775 macaca fasc
146	6	1.8	236	1	YAA3_SCHPO	Q09797 schizosacch	219	6	1.8	316	1	DDL_ENTGA	Q47823 enterococcu
147	6	1.8	239	1	Y247_MYCPN	P75428 mycoplasma	220	6	1.8	318	1	OTC_BACTN	Q8a1e9 bacteroides
148	6	1.8	240	1	BIOD_CLOAB	Q97jc5 clostridium	221	6	1.8	318	1	XERC_LEPIN	O7zam8 leptospira
149	6	1.8	242	1	Y538_VIBPA	O87886 vibrio para	222	6	1.8	322	1	YQBD_BACSU	P45920 bacillus su
150	6	1.8	243	1	GLPF_MYCGA	P52280 mycoplasma	223	6	1.8	323	1	MC3R_MOUSE	P33033 mus musculu
151	6	1.8	250	1	YGB7_MYCPN	Q50315 mycoplasma	224	6	1.8	323	1	MC3R_RAT	P32244 rattus norv
152	6	1.8	251	1	ATP6_MYCLE	P45829 mycobacteri	225	6	1.8	323	1	RAF_MSV36	P00532 murine sarc
153	6	1.8	255	1	TPIS_COXBU	Q83bq3 coxiella bu	226	6	1.8	325	1	ATPT_YEAST	P22135 saccharomyc
154	6	1.8	256	1	COAT_ICMW	O8583 indian cass	227	6	1.8	326	1	LACD_LACLA	P26593 lactococcus
155	6	1.8	256	1	PCNA_NPVAC	P11038 autographa	228	6	1.8	326	1	LACD_STRPN	Q97q13 streptococ
156	6	1.8	258	1	Y789_TREPA	O83768 treponema p	229	6	1.8	326	1	LACD_STRPN	Q8dp2 streptococ
157	6	1.8	260	1	PTB2_ANASP	Q8y52 anabaena sp	230	6	1.8	326	1	VS09_ROTRO	Q91pd4 equine rota
158	6	1.8	261	1	THIG_MYCLE	Q9zb12 mycobacteri	231	6	1.8	326	1	Y067_CHLTR	Q9s529 chlamydia t
159	6	1.8	263	1	LEF5_NPVOP	O10344 orgyia pseu	232	6	1.8	326	1	YN29_CAEEL	P45969 caenorhabdi
160	6	1.8	264	1	RPOD_METTH	O26144 methanobact	233	6	1.8	327	1	ALX_STRPU	Q26657 strongyloce
161	6	1.8	264	1	YJ15_SCHPO	O13679 schizosacch	234	6	1.8	328	1	RPOA_WIGBR	Q8diy8 wigglewort
162	6	1.8	265	1	3BH2_MOUSE	P26149 m 3 beta-hy	235	6	1.8	330	1	GLK_EACHD	Q9kcz4 bacillus ha
163	6	1.8	266	1	MCR2_METJA	Q60367 methanococ	236	6	1.8	331	1	G3P2_KLUJA	Q01077 kluyveromyc
164	6	1.8	267	1	FSD_HELPY	O25911 helicobacte	237	6	1.8	331	1	LDHD_TREPA	O83080 trepanema p
165	6	1.8	267	1	SUHB_PASMU	Q9cnv8 pasteurella	238	6	1.8	333	1	SYW_UREPA	Q9pqw8 ureaplasma
166	6	1.8	268	1	143C_ARATH	Q9c5w6 arabidopsis	239	6	1.8	334	1	FCN1_MOUSE	O70165 mus musculu
167	6	1.8	269	1	AAKB_RAT	P80386 rattus norv	240	6	1.8	334	1	G3P_FICCI	Q9uvc0 pichia cife
168	6	1.8	269	1	VG51_HSVSA	Q01036 herpesvirus	241	6	1.8	334	1	ILVC_METJA	O58938 methanococ
169	6	1.8	274	1	Y440_MYCGE	P47678 mycoplasma	242	6	1.8	335	1	DBX1_MOUSE	P52950 mus musculu
170	6	1.8	274	1	Y44B_MYCPN	P75154 mycoplasma	243	6	1.8	335	1	FCN1_RAT	Q9wt58 rattus norv
171	6	1.8	274	1	YFDQ_ECOLI	P76513 escherichia	244	6	1.8	335	1	KIME_PYRAB	Q9v187 pyrococcus
172	6	1.8	277	1	Y44A_MYCPN	P75151 mycoplasma	245	6	1.8	337	1	TALI_HUMAN	P37837 homo sapien
173	6	1.8	281	1	PANC_CLOAB	Q97f38 clostridium	246	6	1.8	337	1	TALI_MOUSE	Q93092 mus musculu
174	6	1.8	282	1	STR_STAUA	P12055 staphylococ	247	6	1.8	337	1	TALI_RAT	Q9eq80 rattus norv
175	6	1.8	282	1	YC80_GUTH	O78449 guillardia	248	6	1.8	337	1	YXAG_BACSU	P42106 bacillus su
176	6	1.8	283	1	AROE_METTH	O26344 methanobact	249	6	1.8	338	1	HRCA_STRCO	O52163 streptomyce
177	6	1.8	285	1	AROE_CAUCR	Q9ac57 caulobacter	250	6	1.8	338	1	G3P_WASBA	Q9rd6 streptomyce
178	6	1.8	287	1	HUS1_SCHPO	P78955 schizosacch	251	6	1.8	339	1	NCF4_HUMAN	Q9n655 mastigamoeb
179	6	1.8	288	1	SUCD_THETH	P09143 thermus the	252	6	1.8	339	1		Q15080 homo sapien

253 1.8 339 1 NCF4 MOUSE  
254 1.8 340 1 TRPD AQUAE  
255 1.8 341 1 YMG9 VIBU  
256 1.8 342 1 BMAP HUMAN  
257 1.8 343 1 RTCA PYRAB  
258 1.8 344 1 NAGZ YERPE  
259 1.8 345 1 YQ56 METAC  
260 1.8 346 1 YQ56 METAC  
261 1.8 347 1 YQ56 METAC  
262 1.8 348 1 FTSY MYCPN  
263 1.8 349 1 MO2L ARATH  
264 1.8 350 1 EUTR ECOLI  
265 1.8 351 1 PANE SCHPO  
266 1.8 352 1 KITH HSWVD  
267 1.8 353 1 YPBB BAGSU  
268 1.8 354 1 DHAS METJA  
269 1.8 355 1 E2B XYLEA  
270 1.8 356 1 WN11 BRARE  
271 1.8 357 1 CC10 CANAL  
272 1.8 358 1 FIBP CERAE  
273 1.8 359 1 FIBP CERAE  
274 1.8 360 1 DDL ENTHR  
275 1.8 361 1 RFL CORGL  
276 1.8 362 1 AROC VIBAN  
277 1.8 363 1 AROC VIBPA  
278 1.8 364 1 ID12 AERPE  
279 1.8 365 1 V227 FOWPV  
280 1.8 366 1 FIBP HUMAN  
281 1.8 367 1 YGDE ECOLI  
282 1.8 368 1 YQ24 MYCGE  
283 1.8 369 1 DAG2 ARATH  
284 1.8 370 1 YF00 SYN3  
285 1.8 371 1 TRMU MYCPE  
286 1.8 372 1 CL46 BOVIN  
287 1.8 373 1 MNCP OXTR  
288 1.8 374 1 YB2X HAEIN  
289 1.8 375 1 3BH1 MESAU  
290 1.8 376 1 3BH2 MESAU  
291 1.8 377 1 3BH3 MESAU  
292 1.8 378 1 3BH3 MOUSE  
293 1.8 379 1 3BH4 MOUSE  
294 1.8 380 1 PEX3 MOUSE  
295 1.8 381 1 PEX3 RAT  
296 1.8 382 1 PEX3 CRIL  
297 1.8 383 1 PEX3 HUMAN  
298 1.8 384 1 SYW AERPE  
299 1.8 385 1 SYW PYRAE  
300 1.8 386 1 CYB AMBCI

## ALIGNMENTS

RESULT 1  
MO2L HUMAN STANDARD; PRT; 334 AA.  
AC Q9H9S4; Q9BZ33;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE MO25-like protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 4-334 FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,

RA Ninomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 276-334 FROM N.A.  
RA Pearce A.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Mo25 family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AK022639; BAB14147.1; ALT\_INIT.  
CC EMBL; AL138875; CAC28084.1; -  
CC InterPro; IPR008938; ARM.  
CC InterPro; IPR004892; MO25.  
CC Pfam; PF03204; MO25; 1.  
SQ SEQUENCE 334 AA; 38728 MW; 97702273D8548432 CRC64;

Query Match 69.1%; Score 233; DB 1; Length 334;  
Best Local Similarity 99.7%; Pred. No. 5.5e-237;  
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 MFLFSKSHRNPAIVKILNDNLALAEKQDKTKDKASEEVSQSLQAMKEILCGTNEKEPPT 63  
DB 1 MFLFSKSHRNPAIVKILNDNLALAEKQDKTKDKASEEVSQSLQAMKEILCGTNEKEPPT 60  
QY 64 EAVAQAELYSGLLVTLADIQLIDFGKQDVTOIFNNILRRQIGTRSPVVEYISAHP 123  
DB 61 EAVAQAELYSGLLVTLADIQLIDFGKQDVTOIFNNILRRQIGTRSPVVEYISAHP 120  
QY 124 HILFMLKGVAPQIALRGIMLRECIHEPLAKITLFSNQDFEFKYVELSTFDIASDA 183  
DB 121 HILFMLKGVAPQIALRGIMLRECIHEPLAKITLFSNQDFEFKYVELSTFDIASDA 180  
QY 184 FATFKDLLTRHKVLVADFLQNYDTTFEDYEKLQSENVTYKRSQSLKGLGELLDRHFA 243  
DB 181 FATFKDLLTRHKVLVADFLQNYDTTFEDYEKLQSENVTYKRSQSLKGLGELLDRHFA 240  
QY 244 IMTKYISKPENLKLMMNLDRKSPNQFAHVFVKVFAVSPHKTQPIVEILLKNQPKLIE 303  
DB 241 IMTKYISKPENLKLMMNLDRKSPNQFAHVFVKVFAVSPHKTQPIVEILLKNQPKLIE 300  
QY 304 FLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 337  
DB 301 FLSSFOKERTDDEQFADEKNYLIKQIRDLKKTAP 334

## RESULT 2

MO2L MOUSE STANDARD; PRT; 334 AA.  
AC Q9DBI6; Q8BG52; Q91WB8; Q91YL0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE MO25-like protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC STRAIN=C57BL/6J;  
RC TISSUE=Cerebellum, Eye, Pituitary, and Testis;  
RX MEDLINE=22354683; PubMed=12466851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schirnd L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brucic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,  
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Karyaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,  
RA Verardo L., Wagner L., Wallesstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573 (2002).  
RN [2].  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=FVB/N; TISSUE=Mammary gland, and Salivary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9DB16-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9DB16-2; Sequence=VSP 007417, VSP 007418;  
CC Note=No experimental confirmation available;  
CC -!- SIMILARITY: Belongs to the MO25 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AK005323; BAB23953.2; ALT\_INIT.  
CC EMBL; AK030474; BAC26978.1; ALT\_INIT.  
CC EMBL; AK053642; BAC35457.1; ALT\_INIT.  
CC EMBL; AK076758; BAC36470.1; ALT\_INIT.  
CC EMBL; AK076867; BAC36513.1; -.  
CC EMBL; BC016128; AAH16128.1; -.

DR EMBL; BC016546; AAH16546.1; -.  
DR MGD; MGI:1916258; 1500031K13Rik.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR004892; MO25.  
DR Pfam; PF03204; MO25; 1.  
KW Alternative splicing  
FT VARSPLIC 276 293 VFVSPHKTPQIVEILLK -> NSVFITRIHKLKRWLSS  
FT (in isoform 2).  
FT /FTID=VSP 007417.  
FT VARSPLIC 294 334 Missing (in isoform 2).  
FT /FTID=VSP 007418.  
FT CONFLICT 42 42 S -> P (IN REF. 1; BAB23953).  
FT CONFLICT 229 239 L -> R (IN REF. 2; AAH16546).  
SQ SEQUENCE 334 AA; 38718 MW; 222F04A87FB4EB6F CRC64;  
  
Query Match 35.9%; Score 121; DB 1; Length 334;  
Best Local Similarity 100.0%; Pred. No. 3.8e-119;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 122 HPFILMLKGYEAPQIALRCGIMLRECIHHEPLAKIILFNSQFRDFFKVELSTFDIAS 181  
Db 119 HPFILMLKGYEAPQIALRCGIMLRECIHHEPLAKIILFNSQFRDFFKVELSTFDIAS 178  
Qy 182 DAFATFKDLLTRHKVLVADFLQNYDTIPEDYEKLQSENYVTKQSLKGLGLILDRHN 241  
Db 179 DAFATFKDLLTRHKVLVADFLQNYDTIPEDYEKLQSENYVTKQSLKGLGLILDRHN 238  
Qy 242 F 242  
Db 239 F 239  
  
RESULT 3  
MO25 HUMAN STANDARD; PRT; 341 AA.  
AC Q9Y376;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 18-OCT-2003 (Rel. 42, Last annotation update)  
DE MO25 protein (CGI-66).  
GN MO25.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20272150; PubMed=10810093;  
RA Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;  
RT "Identification of novel human genes evolutionarily conserved in  
RT Caenorhabditis elegans by comparative proteomics";  
RL Genome Res. 10:703-713 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Hypothalamus;  
RA Jin W., Shi J., Ren S., Gu J., Fu S., Huang Q., Dong H., Yu Y., Fu G.,  
RA Wang Y., Chen Z., Han Z.;  
RT "A novel gene expressed in the human hypothalamus";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ductum;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RL CC -1- SIMILARITY: Belongs to the Mo25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF151824; AAD34061.1; -
DR EMBL; AF113536; AAF14873.1; -
DR EMBL; BC020570; AAH20570.1; -
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39869 MW; EC710A528B659811 CRC64;

Query Match 6.8%; Score 23; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.9e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266
Db 245 IMTKYISKPENLKMNNLLRDKS 267

RESULT 4
MO25 MOUSE
ID _MO25_MOUSE STANDARD; PRT; 341 AA.
AC Q06138;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mo25 protein.
GN MO25 OR CAB39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93119656; PubMed=8418809;
RA Miyamoto H., Matsushiro A., Nozaki M.;
RT "Molecular cloning of a novel mRNA sequence expressed in cleavage
RT stage mouse embryos."
RL Mol. Reprod. Dev. 34:1-7(1993).
CC -1- FUNCTION: ONE OF THE FIRST GENES TO BE TRANSCRIBED DURING MOUSE
CC DEVELOPMENT, MAY PLAY SOME GENERAL FUNCTION.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- DEVELOPMENTAL STAGE: TRANSCRIBED DURING EARLY MOUSE DEVELOPMENT.
CC DETECTED AT ALL DEVELOPMENTAL STAGES FROM THE EGG THROUGH THE
CC BLASTOCYST, MOST ABUNDANT AT THE 2-CELL STAGE.
CC -1- SIMILARITY: Belongs to the Mo25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; S51858; AAB24801.1; -.
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39842 MW; E7F668529D6FE811 CRC64;

Query Match 6.8%; Score 23; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.9e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266
Db 245 IMTKYISKPENLKMNNLLRDKS 267

RESULT 5
DE76 CHLPR
ID _DE76_CHLPR STANDARD; PRT; 321 AA.
AC Q9XFY6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Degreening related gene dee76 protein.
GN DEE76.
OS Chlorella protothecoides.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Auxenochlorella.
OX NCBI_TaxID=3075;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ACC25;
RX MEDLINE=20256472; PubMed=10798614;
RA Hertensteiner S., Chinner J., Matile P., Thomas H., Donnison I.S.;
RT "Chlorophyll breakdown in Chlorella protothecoides: characterization
RT of degreening and cloning of degreening-related genes."
RL Plant Mol. Biol. 42:439-450(2000).
CC -1- SIMILARITY: Belongs to the Mo25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ238632; CAB42595.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 321 AA; 37262 MW; 918FD02964B09071 CRC64;

Query Match 4.7%; Score 16; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 ASDAFATFKDLLTRHK 195
Db 168 ASDAFATFKDLLTRHK 183

RESULT 6
YFV6 SCHPO
ID _YFV6_SCHPO STANDARD; PRT; 329 AA.
AC Q9P7Q8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C1834.06c in chromosome I.
GN SPAC1834.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
```

```

OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scourto J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA McInerney P., Moulle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmons M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
RA Waltjens I., Vanstreels E., Rieger M., Schaefer W., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
RA Dega R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipakovski G.V., Ussery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL157734; CAB75774.1; -
DR PIR; T50117; T50117.
DR GeneDB_Spombe; SPAC1834.06c; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
KW Hypothetical protein.
SQ SEQUENCE 329 AA; 38521 MW; 073DD0607A64C952 CRC64;

Query Match 4.7%; Score 16; DB 1; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 219 SENVTVTKRQSLKGLGE 234
Db 216 SENVTVTKRQSLKGLGE 231

RESULT 7
Mo25_DROME
ID Mo25_DROME STANDARD; PRT; 339 AA.
AC P91891; Q9VW85;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mo25 protein (dMo25).
GN Mo25 OR C54083.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

```

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov E.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosler C., Gabrielian A.B., Garg N.S., Galbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpun G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-F., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000402; BAA19098.1; -
DR EMBL; AE003526; BAF49432.1; -
DR FlyBase; FBgn0017572; Mo25.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
FT CONFLICT 51 51 Y -> H (IN REF. 1).
FT CONFLICT 102 102 V -> L (IN REF. 1).
SQ SEQUENCE 339 AA; 39385 MW; 5790BD91754C1C74 CRC64;

Query Match 4.5%; Score 15; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;

```



Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 105 LRRQIGTRSPTEVII 119  
 Db 103 LRRQIGTRSPTEVII 117

RESULT 8  
 ID MO2M CAEEL STANDARD; PRT; 338 AA.  
 AC O18211; (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Hypothetical MO25-like protein Y53C12A.4 in chromosome II.  
 DN Y53C12A.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Kershaw J., Lennard N.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the Mo25 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ001157; CAA04556.1; -;  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 SQ SEQUENCE 384 AA; 44392 MW; 1D0C34A35D9116F5 CRC64;

Query Match 4.2%; Score 14; DB 1; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 105 LRRQIGTRSPTEVII 118  
 Db 109 LRRQIGTRSPTEVII 122

RESULT 9  
 ID HYMA EMENI STANDARD; PRT; 384 AA.  
 AC O60032;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DE Hypothetical MO25-like protein Y53C12A.4 in chromosome II.  
 DN Y53C12A.4.  
 OS Conidiophore development protein hYMA.  
 OC Hymenochaeta; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 NCBI\_TaxID=162425;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC MEDLINE=99126010; PubMed=9928930;  
 RA Karos M., Fischer R.;  
 RL "Molecular characterization of HymA, an evolutionarily highly  
 conserved and highly expressed protein of Aspergillus nidulans.";  
 Mol. Gen. Genet. 260:510-521(1999).

CC -!- FUNCTION: Required for conidiophore development.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the Mo25 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ001157; CAA04556.1; -;  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 SQ SEQUENCE 384 AA; 44392 MW; 2E203D0DD110C5FD6 CRC64;

Query Match 2.7%; Score 9; DB 1; Length 384;  
 Best Local Similarity 100.0%; Pred. No. 0.29;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 274 FHVKVFEVA 282  
 Db 301 FHVKVFEVA 309

RESULT 10  
 ID TOR DROME STANDARD; PRT; 923 AA.  
 AC P18475;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DE Tyrosine-protein kinase receptor torso precursor (EC 2.7.1.112).  
 DN TOR.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC MEDLINE=89181943; PubMed=2927509;  
 RA Sprenger F., Stevens L.M., Nusslein-Volhard C.;  
 RL "The Drosophila gene torso encodes a putative receptor tyrosine  
 kinase.";  
 Nature 338:478-483(1989).  
 [2]  
 RN CHARACTERIZATION.  
 RC MEDLINE=93140754; PubMed=8423783;  
 RA Sprenger F., Torsoclar M.M., Morrison D.K.;  
 RL "Biochemical analysis of torso and D-raf during Drosophila  
 embryogenesis: implications for terminal signal transduction.";  
 Mol. Cell. Biol. 13:1163-1172(1993).  
 CC -!- FUNCTION: Probable receptor with tyrosine-protein kinase activity.  
 CC Required for determination of anterior and posterior terminal  
 CC structures in the Drosophila embryo. The ligand of torso seems to  
 CC be TSL.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed throughout the embryo but is  
 CC activated specifically at the poles.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



Salanoubat M., Genin S., Artiguenave F., Gouzy J., Manganot S., Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L., Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N., Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T., Siguiet P., Thebault P., Whalen M., Wincker P., Levy M., Weissenbach J., Boucher C.A.;  
"Genome sequence of the plant pathogen *Ralstonia solanacearum*.";  
RL Nature 415:497-502 (2002).  
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton gradient across the membrane.  
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).  
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) has five core - and CF(0) - the membrane proton channel. CF(1) has five subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0) has three main subunits: a, b and c.  
CC -!- SIMILARITY: Belongs to the ATPase epsilon chain family.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AL646081; CAD17961.1; -.  
CC HAMAP; MF 00530; -; 1.  
CC InterPro; IPR001469; ATPsynth\_DE.  
CC Pfam; PF00403; ATP-synt\_DE; 1.  
CC Pfam; PF02823; ATP-synt\_DE\_N; 1.  
CC ProDom; PD000944; ATPsynth\_DE; 1.  
CC TIGRFAMs; TIGR01216; ATP\_synth\_epsi; 1.  
CC Hydroxylase; ATP01216; Hydrogen ion transport; Plasmid;  
KW Complete proteome.  
KW SEQUENCE 139 AA; 15304 MW; A506B25345D28C1E CRC64;  
SQ  
Query Match 2.1%; Score 7; DB 1; Length 139;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 64 EAAVQLA 70  
DB 123 EAAVQLA 129  
-----  
RESULT 12  
Y003 METJA  
ID Y003 METJA STANDARD; PRT; 156 AA.  
AC Q60313;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0003.  
GN MJ0003.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Adams M.D., Reich C.I., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Merrick J.M., Glodek A., Overbeek R., Kirkness E.F., Weinstock K.G., Weidman J.F., Fuhmann J.L., Nguyen D., Scott J.L., Geoghagen N.S.M., Peterson J.D., Sadow P.W., Hanna M.C., Utterback T.R., Kelley J.M., Hurst M.A., Kaine B.P., Borodovsky M., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA "Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.";  
RT

```

RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67459; AAB97990.1; -.
DR PIR; C64300; C64300.
DR TIGR; M00003; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 156 AA; 18428 MW; 364797DB0E38AB9 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 VEILKN 297
DB 34 VEILKN 40

RESULT 13
ATPD BUCAP STANDARD; PRT; 177 AA.
ID ATPD BUCAP
AC O51875;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase delta chain (EC 3.6.3.14).
GN ATP OR BUGS005
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97361981; PubMed=9216881;
RA Clark M.A., Baumann P.;
RT "The (F1F0) ATP synthase of Buchnera aphidicola (endosymbiont of
RT aphids): genetic analysis of the putative ATP operon.";
RL Curr. Microbiol. 35:84-89(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184963; PubMed=9516544;
RA Clark M.A., Baumann L., Baumann P.;
RT "Sequence analysis of a 34.7-kb DNA segment from the genome of
RT Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,
RT the atp operon, gldA, and rho.";
RL Curr. Microbiol. 36:158-163(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -!- FUNCTION: This protein seems to be part of the stalk that links
CC CF(0) to CF(1). It either transmits conformational changes from
CC CF(0) into CF(1) or is implicated in proton conduction.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(-)(n) = ADP + phosphate +
CC H(+)(Out)
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c.
CC -!- SIMILARITY: Belongs to the ATPase delta chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF008210; AAC38113.1; -.
DR EMBL; AE014076; AAM67577.1; -.
DR HSSP; P00831; IABV.
DR InterPro; IPR000711; ATPynt_OSCP.
DR Pfam; PF00213; OSCP; 1.
DR PRINTS; PR00125; ATPASEDELTA.
DR TIGR; TIGR01145; ATP synt_delta; 1.
DR PROSITE; PS00389; ATPASE DELTA; FALSE NEG.
KW Hydrolase; ATP synthetase; CF(1); Hydrogen ion transport;
KW Complete proteome.
SQ SEQUENCE 177 AA; 20671 MW; E1BD82A7529C01B7 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 IFNNILR 106
DB 88 IFNNILR 94

RESULT 14
GTA3 CHICK STANDARD; PRT; 228 AA.
ID GTA3 CHICK
AC P26697;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutathione S-transferase 3 (EC 2.5.1.18) (GST-CL3) (GST class-alpha).
OS Gallus gallus (Chicken)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=White leghorn; TISSUE=Liver;
RX MEDLINE=92149826; PubMed=1339283;
RA Chang L.-H., Fan J.-Y., Liu L.-F., Tsai S.-P., Tam M.F.;
RT "Cloning and expression of a chick liver glutathione S-transferase CL
RT 3 subunit with the use of a baculovirus expression system.";
RL Biochem. J. 281:545-551(1992).
CC -!- FUNCTION: CATALYZES THE CONJUGATION OF GSH TO A WIDE VARIETY OF
CC ELECTROPHILIC ALKYLATING AGENTS. ALSO INVOLVED IN THE METABOLISM
CC OF LIPID HYDROPEROXIDES, PROSTAGLANDINS AND LEUKOTRIENE A4 AND IN
CC BINDING OF NON-SUBSTRATE HYDROPHOBIC LIGANDS SUCH AS BILE ACIDS, A
CC NUMBER OF DRUGS AND THYROID HORMONES. THIS GST DOES NOT EXHIBIT
CC PEROXIDASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC -!- SUBUNIT: HOMODIMER OR HETERODIMER (WITH A SUBUNIT FROM GROUP
CC CL-4).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: THE VARIATIONS WERE FOUND FROM AA SEQUENCING AND
CC IMPLY THERE ARE MULTIPLE FORMS OF CL-3.
CC -!- SIMILARITY: Belongs to the GST superfamily. Alpha family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38219; AAA62731.1; -.
DR PIR; S19734; S19734.
DR HSSP; P24472; 1GUK.
```

```
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR004045; GST_Nterm.
DR Pfam: PF00043; GST_C; 1.
DR Pfam: PF02798; GST_N; 1.
KW Transferase; MultiGene family.
FT INIT MET 0 0
FT MOD RES 1 1 BLOCKED.
FT VARIANT 44 44 L->V.
FT VARIANT 46 46 S->A.
FT VARIANT 48 48 I->F.
FT VARIANT 48 48 I->V.
FT VARIANT 51 51 F->R.
FT VARIANT 128 129 TS->AN.
FT VARIANT 134 135 AY->VF.
FT VARIANT 154 154 W->R.
FT VARIANT 157 158 IH->VV.
FT VARIANT 162 162 A->T.
FT VARIANT 165 165 M->A.
FT VARIANT 167 167 E->V.
SQ SEQUENCE 228 AA; 26194 MW; C7450E146F41E787 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 YEKLLQS 219
DB 40 YEKLLQS 46

RESULT 15
RADC FUSNN STANDARD; PRT; 232 AA.
AC QGRFT5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE DNA repair protein radC homolog.
GN RADC OR FN0909.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21866394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010599; AL95105.1; -.
CC HANAP; MF_00018; -.
CC InterPro: IPR001405; RadC.
CC Pfam; PF04002; RadC; 1.
CC ProDom; PD007415; RadC; 1.
CC TIGRfam; TIGR00608; radc; 1.
CC PROSITE; PS01302; RADC; FALSE_NEG.
KW DNA repair; Complete proteome.
SQ SEQUENCE 232 AA; 26730 MW; 0FD3D4CE84E45A93 CRC64;
```

```
Query Match 2.1%; Score 7; DB 1; Length 232;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLLGEL 235
DB 83 LKLLGEL 89

RESULT 16
MHPC ECOLI STANDARD; PRT; 293 AA.
AC P77044; P71204; P77205;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-hydroxy-6-ketono-2,4-dienedioic acid hydrolase (EC 3.7.1.-).
GN MHPC OR B0349.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Kawamukai M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / CS520;
RA Ferrandez A., Garcia J.L., Diaz E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -! PAIRWAY: 3-hydroxyphenylpropionate degradation.
CC -! SIMILARITY: STRONG, TO B.CEPACIA AND PSEUDOMONAS BPHD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D86239; BA13054.1; -.
CC EMBL; Y09555; CAAT70749.1; -.
CC EMBL; AE000142; AAC73452.1; ALT_INIT.
CC EMBL; U73857; AAB18073.1; ALT_INIT.
CC MEROPS; S33.UNW; -.
CC EcoGene; EG20275; mbpC.
CC InterPro: IPR000073; A/b hydrolase.
CC InterPro: IPR003089; AB_hydrolase.
CC InterPro: IPR000639; Epox_hydrolase.
CC InterPro: IPR000379; Ser_estrs.
CC Pfam; PF00561; abhydrolase; 1.
CC PRINTS; PR00111; ABHYDROLASE.
```

DR PRINTS; PR00412; EPOXHYDROLASE.  
 KW Aromatic hydrocarbons catabolism; Hydrolase; Complete proteome.  
 FT ACT SITE 90 BY SIMILARITY.  
 FT CONFLICT 158 E -> G (IN REF. 1 AND 2).  
 SQ SEQUENCE 293 AA; 32585 MW; 4407DF7B90EA0E80 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 253 ENLKMM 259  
 Db 173 ENLKMM 179  
 |||||

## RESULT 17

RMAR\_CANGA STANDARD; PRT; 339 AA.  
 AC P21358;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Mitochondrial ribosomal protein VAR1.  
 GN VAR1.  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5478;

[1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=86011564; PubMed=3900417;  
 RA Ainley W.M., Macreadie I.G., Butow R.A.;  
 RT "VAR1 Gene on the mitochondrial genome of Torulopsis glabrata."  
 RL J. Mol. Biol. 184:565-576(1985).  
 CC -!- FUNCTION: Essential for mitochondrial protein synthesis and  
 CC required for the maturation of small ribosomal subunits (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -!- SIMILARITY: BELONGS TO THE VAR1 FAMILY OF RIBOSOMAL PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; X02893; CAA26652.1; -;  
 DR PIR; S04682;  
 DR InterPro; IPR007980; Yeast VAR1.  
 DR Pfam; PF05316; Yeast VAR1; 1.  
 KW Ribosomal protein; Mitochondrion.  
 SQ SEQUENCE 339 AA; 40828 MW; 354820A9C1FE9E50 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYIS 250  
 Db 180 IMTKYIS 186  
 |||||

## RESULT 18

CARP\_POLTU STANDARD; PRT; 340 AA.  
 AC P17576;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Polyporopepsin (EC 3.4.23.29) (Aspartic proteinase).

OS Polyporus tuliferus (Irpe lacteus).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Aphyllophorales; Polyporaceae; Polyporus.  
 OX NCBI\_TaxID=29885;  
 RN SEQUENCE FROM N.A.; AND SEQUENCE OF 1-24.

RA Kobayashi H., Sekibata S., Shibuya H., Yoshida S., Kusakabe I.,  
 RA Murakami K.;  
 RT "Cloning and sequence analysis of cDNA for Irpe lacteus aspartic  
 RT proteinase."  
 RL Agric. Biol. Chem. 53:1927-1933(1989).  
 CC -!- CATALYTIC ACTIVITY: Milk clotting activity, broad specificity, but  
 CC fails to cleave 15-Leu-|-Tyr-16 or 16-Tyr-|-Leu-17 of insulin B  
 CC chain.  
 CC -!- SIMILARITY: Belongs to peptidase family A1.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; D00589; BAA00467.1; -;  
 DR PIR; JU0057; PEIKL.  
 DR HSP; P32329; 1YPS.  
 DR MEROPS; A01.019; -;  
 DR InterPro; IPR001969; Asparticase AS.  
 DR InterPro; IPR009007; Pept A acid;  
 DR InterPro; IPR001461; Peptidase\_A1.  
 DR Pfam; PF00026; asp; 1.  
 DR PRINTS; PR00792; PEPSIN.  
 DR PROSITE; PS00141; ASP PROTEASE; 2.  
 KW Hydrolase; Aspartyl protease; Glycoprotein.  
 FT ACT SITE 32 32  
 FT ACT SITE 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 192 192  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 340 AA; 35050 MW; 9BAF837264D42FEF CRC64;

Query Match 2.1%; Score 7; DB 1; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 IASDAFA 185  
 Db 220 IASDAFA 226  
 |||||

## RESULT 19

MOZM\_ARATH STANDARD; PRT; 343 AA.  
 AC Q9M0M4; O23570;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical MO25-like protein At4g17270.  
 GN AT4G17270 OR DI4670W.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.

OX NCBI\_TaxID=3702;  
 RN SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;  
 RC MEDLINE=98121113; PubMed=9461215;  
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,  
 RA Bergkamp R., Dirke W., van Staveren M., Stiekema W., Drost L.,  
 RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,  
 RA Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T., Terryn N.,  
 RA Gielen J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,

RA Aubourg S., Gy L., Kreis M., Lao N., Kavanagh T., Hempel S.,  
RA Kotter P., Entian K.-D., Rieger M., Schaefer M., Funk B.,  
RA Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,  
RA Puigdomenech P., Douka A., Voukelatou E., Milioni D., Hatzopoulos P.,  
RA Piravandi E., Obermaier B., Hilbert H., Duesterhoeft A., Moeres T.,  
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansoerge W.,  
RA Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H.-W.,  
RA Klosterman S., Schueller C., Chalwatzi N.,  
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
RT Arabidopsis thaliana."  
RL Nature 391:485-488(1998).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20093488; PubMed=10617198;  
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
RA Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M.,  
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,  
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,  
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,  
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,  
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,  
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
RA Mocijman F., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
RA De Keyser A., Buyschaert C., Gielens J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,  
RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,  
RA Neumann S., Argirou A., Vitale B., Liguori R., Piravandi E.,  
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Chedfor E., Cooke R., Berger C., Montfort A., Casacuberta E.,  
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,  
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Bernhoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidan M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana."  
RL Nature 402:769-777(1999).  
[3]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=22954850; PubMed=14593172;  
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
RA Arakawa T., Bann J., Sanno F., Bowser L., Brooks S.Y., Carninci P.,  
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
RA Hayaashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,

RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Empirical analysis of transcriptional activity in the Arabidopsis  
RT genome."  
RL Science 302:842-846(2003).  
CC -!- SIMILARITY: Belongs to the Mo25 family.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous  
CC gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z97343; CAB10508.1; ALT\_SEQ.  
DR EMBL; AL161546; CAB78730.1; -.  
DR EMBL; AF380659; AAK55740.1; -.  
DR InterPro: IPR008938; ARM.  
DR InterPro: IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 343 AA; 39650 MW; D340B49A4924B7D1 CRC64;  
  
Query Match 2.1%; Score 7; DB 1; Length 343;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 272 EAFHVK 278  
Db 278 EAFHVK 284  
|||||  
-----  
RESULT 20  
MO2N ARATH STANDARD; PRT; 343 AA.  
AC Q9FGK3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical MO25-like protein At5g47540.  
GN AT5G47540 OR MNJ7.13.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the Mo25 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB025628; BAB09080.1; -.  
DR InterPro: IPR008938; ARM.  
DR InterPro: IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
KW Hypothetical protein.

```
SQ SEQUENCE 343 AA; 39457 MW; 46950D6A9A82FBB5 CRC64;
Query Match 2.1%; Score 7; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 272 EAFHVK 278
DB 278 EAFHVK 284
|||||
RESULT 21
OCD BRUSU STANDARD; PRT; 358 AA.
ID Q8FVE4;
AC 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ornithine cyclodeaminase (EC 4.3.1.12) (OCD).
GN ARCB OR BRA0899.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCB1_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kralj B., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tetzelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
CC -!- CATALYTIC ACTIVITY: L-ornithine = L-proline + NH(3).
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Arginine degradation; second step.
CC -!- SIMILARITY: Belongs to the ornithine cyclodeaminase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U57319; AAC05589.1; -
DR InterPro; IPR003462; ODC_Mu_crystal1.
DR Pfam; PF02423; ODC_Mu_crystal1.1.
KW Lyase; NAD; Arginine metabolism.
SQ SEQUENCE 359 AA; 39420 MW; 94CFD9C69CB785F4 CRC64;
Query Match 2.1%; Score 7; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 214 EKLLQSE 220
DB 347 EKLLQSE 353
|||||
RESULT 23
OCD BRUME STANDARD; PRT; 359 AA.
ID -OCD BRUME STANDARD; PRT; 359 AA.
AC Q8YCV1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ornithine cyclodeaminase (EC 4.3.1.12) (OCD).
GN ARCB OR BME10397.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCB1_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muej C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Gotsman E.,
RA Seikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
CC -!- CATALYTIC ACTIVITY: L-ornithine = L-proline + NH(3).
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Arginine degradation; second step.
CC -!- SIMILARITY: Belongs to the ornithine cyclodeaminase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE014583; AAN34071.1; -
DR TIGR; BRA0899; -
DR InterPro; IPR003462; ODC_Mu_crystal1.
DR Pfam; PF02423; ODC_Mu_crystal1.1.
KW Lyase; NAD; Arginine metabolism; Complete proteome.
SQ SEQUENCE 358 AA; 39299 MW; DA2F02B1171B4C81 CRC64;
Query Match 2.1%; Score 7; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 214 EKLLQSE 220
DB 346 EKLLQSE 352
|||||
RESULT 22
OCD BRUAB STANDARD; PRT; 359 AA.
ID Q89I75;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE009677; AAL53639.1; --  
DR PIR; AD3559; AD3559; ODC\_Mu\_crystall.  
DR InterPro; IPR003462; ODC\_Mu\_crystall.  
DR Pfam; PF02423; ODC\_Mu\_crystall; 1.  
KW Lyase; NAD; Arginine metabolism; Complete proteome.  
SQ SEQUENCE 359 AA; 39460 MW; F29E90021EF950B5 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EKLQSE 220  
Db 347 EKLQSE 353

RESULT 24  
P65H MYCGE STANDARD; PRT; 372 AA.  
AC P47459;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Proline-rich P65 protein homolog.  
GN MG217.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes;  
OX NCBI\_TaxID=2097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
RA Nguyen D.T., Uutterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.-F., Dougherty B.A., Bitt K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,  
RT "The minimal gene complement of Mycoplasma genitalium."  
RL Science 270:397-403 (1995).  
CC -1- SIMILARITY: TO M. PNEUMONIAE PROTEIN P65.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U39701; AAC71436.1; --  
DR PIR; I64223; I64223.  
DR TIGR; MG217; --  
KW Repeat; Complete proteome.  
FT DOMAIN 29 159 8 X DPNQ (Q) FNQ REPEATS.  
FT REPEAT 29 40  
FT REPEAT 41 52  
FT REPEAT 53 60  
FT REPEAT 61 72  
FT REPEAT 73 80  
FT REPEAT 81 92  
FT REPEAT 93 100  
FT REPEAT 101 112  
FT REPEAT 113 119  
FT REPEAT 120 131  
FT REPEAT 132 138

FT REPEAT 139 150 1-7.  
FT REPEAT 151 162 1-8.  
SQ SEQUENCE 372 AA; 44664 MW; 4C29701D213CE19E CRC64;

Query Match 2.1%; Score 7; DB 1; Length 372;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMNLLR 263  
Db 324 LMNLLR 330

RESULT 25  
BIOF AQUAE STANDARD; PRT; 373 AA.  
AC O66875;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 8-amino-7-oxononanoate synthase (EC 2.3.1.47) (AONS) (8-amino-7-  
DE ketopelargonate synthase) (7-keto-8-amino-pelargonate synthetase)  
DE (7-KAP synthetase) (L-alanine-pimelyl CoA ligase).  
GN BIOF OR AQ 626.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE=98196666; PubMed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus."  
RL Nature 392:353-358 (1998).  
CC -1- CATALYTIC ACTIVITY: 6-carboxyhexanoyl-CoA + L-alanine = 8-amino-7-  
CC oxononanoate + CoA + CO(2).  
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).  
CC -1- PATHWAY: Biotin biosynthesis; first step.  
CC -1- SIMILARITY: Belongs to class-II of pyridoxal-phosphate-dependent  
CC aminotransferases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AR000699; AAC06836.1; --  
DR PIR; G70355; G70355.  
DR HSSP; P12998; IBSO.  
DR InterPro; IPR003408; Ala\_synthase.  
DR InterPro; IPR004839; Aminotrans\_I/II.  
DR InterPro; IPR001917; Aminotrans\_II.  
DR InterPro; IPR004723; BioF.  
DR Pfam; PF02490; ALA synthase; 1.  
DR Pfam; PF00155; aminotran\_1\_2; 1.  
DR TIGR; TIGR00858; bioF; 1.  
DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
KW Biotin biosynthesis; Transferrase; Pyridoxal phosphate;  
KW Complete proteome.  
FT BINDING 223 223 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 373 AA; 42532 MW; 536B34A5D5F84401 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 373;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305



Db 274 PKLIEPL 280  
|||||  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 26  
HYMI YEAST  
ID HYMI YEAST STANDARD; PRT; 399 AA.  
AC P32454;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE HYMI protein.  
GN HYMI OR YKL189W.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GRF88;  
RA MEDLINE=93348778; PubMed=8394042;  
RA Cheret G., Mattheakis L.C., Sor F.;  
RT "DNA sequence analysis of the YCN2 region of chromosome XI in  
Saccharomyces cerevisiae.";  
RL Yeast 9:661-667(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=94205264; PubMed=8154185;  
RA Wiemann S., Voss H., Schwager C., Rupp T., Stegemann J.,  
RA Zimmermann J., Grothues D., Sengen C., Erfle H., Hewitt N.,  
RA Banrevi A., Ansorge W.;  
RT "Sequencing and analysis of 51.6 kilobases on the left arm of  
chromosome XI from Saccharomyces cerevisiae reveals 23 open reading  
frames including the FAS1 gene.";  
RL Yeast 9:1343-1348(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Maia e Silva A., Bossier P., Vilela C., Fernandes L., Soares H.,  
RA Guerreiro P., Rodrigues-Pousada C.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RP GENE NAME.  
RX MEDLINE=20157038; PubMed=10655212;  
RA Dorland S., Deegenars M.L., Stillman D.J.;  
RT "Roles for the Saccharomyces cerevisiae SDS3, CBK1 and HYMI genes in  
transcriptional repression by SIN3.";  
RL Genetics 154:573-586(2000).  
CC -!- SIMILARITY: Belongs to the Mo25 family.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X69765; CAA49422.1; -.  
DR EMBL; X74151; CAA52249.1; -.  
DR EMBL; Z28189; CAA82032.1; -.  
DR PIR; S34681; S34681.  
DR GeneOnline; 139944; -.  
DR SGD; S0001672; HYMI.  
DR GO; GO:0005622; C:intracellular; IDA.  
DR GO; GO:0016564; P:transcriptional repressor activity; IMP.  
DR GO; GO:0007109; P:cytokinesis, completion of separation; IMP.  
DR GO; GO:0008360; P:regulation of cell shape; IGI.  
DR InterPro; IPR008938; ARM.  
DR Pfam; PF03204; Mo25; 1.  
SQ SEQUENCE 399 AA; 45853 MW; F48860754C892BA9 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 399;

Oy 222 YVTKRQS 228  
|||||  
Db 244 YVTKRQS 250  
|||||

Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 27  
EXOF RHIME  
ID EXOF RHIME STANDARD; PRT; 421 AA.  
AC Q02728;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Exopolysaccharide production protein exoF precursor.  
EX EXOF OR RB1068 OR SMB20945.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OG Plasmid pSymB (megaplasmid 2).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RCR2011 / SU47;  
RX MEDLINE=93177026; PubMed=8439670;  
RA Mueller P., Keller M., Weng W.M., Quandt J., Arnold W., Puehler A.;  
RT "Genetic analysis of the Rhizobium meliloti exoYFQ operon: ExoY is  
homologous to sugar transferases and ExoQ represents a transmembrane  
protein.";  
RL Mol. Plant Microbe Interact. 6:55-65(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396508; PubMed=11481431;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,  
RA Golding B., Puehler A.;  
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-  
fixing endosymbiont Sinorhizobium meliloti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
CC -!- FUNCTION: INVOLVED IN SUCCINOGLYCAN (EPS 1) SYNTHESIS. NEEDED FOR  
THE ADDITION OF THE FIRST SUGAR (GALACTOSE) TO THE ISOPRENOL  
CARRIER.  
CC -!- PATHWAY: Exopolysaccharide biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L05588; AAA26265.1; -.  
DR EMBL; AL603645; CAC49468.1; -.  
DR PIR; D95975; D95975.  
DR InterPro; IPR003715; Poly export.  
DR Pfam; PF02563; Poly export; 1.  
DR Exopolysaccharide synthase; Plasmid; Periplasmic; Signal;  
KW Complete proteome.  
FT SIGNAL 1 31  
FT CHAIN 32 421  
FT POTENTIAL.  
SQ SEQUENCE 421 AA; 45871 MW; 55D8EC55B3A469EF CRC64;  
EXOF.

Query Match 2.1%; Score 7; DB 1; Length 421;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 67 AQLAQEL 73  
|||||

```
Db          336 AQAQEL 342

RESULT 28
P65_MYCPN
ID   P65_MYCPN          STANDARD;          PRT;   423 AA.
AC   P53663; P53664;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Proline-rich P65 protein.
GS   P65 OR MPN309 OR MP527.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=95286529; PubMed=7768845;
RA   Proft T., Hilbert H., Layh-Schmitt G., Herrmann R.;
RT   "The proline-rich P65 protein of Mycoplasma pneumoniae is a component
RT   of the Triton X-100-insoluble fraction and exhibits size polymorphism
RT   in the strains M129 and FH.";
RL   J. Bacteriol. 177:3370-3378(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97252497; PubMed=9098066;
RA   Krause D.C., Proft T., Hedreya C.T., Hilbert H., Plagens H.,
RA   Herrmann R.;
RT   "Transposon mutagenesis reinforces the correlation between Mycoplasma
RT   pneumoniae cytoskeletal protein HMW2 and cytoadherence.";
RL   J. Bacteriol. 179:2668-2677(1997).
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED, WITH EPITOPES EXPOSED
CC   AT THE CELL SURFACE (PROBABLE).
CC   -!- DOMAIN: THE PENTA/HEXAPEPTIDES REPEATS FORM A PROLINE-RICH ACIDIC
CC   DOMAIN. IN ADDITION, A PART OF THIS REGION CONTAINS A PERFECT
CC   DIRECT REPEAT.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED (PROBABLE).
CC   -!- POLYMORPHISM: Duplication in FH strain leads to a proline-rich
CC   domain composed of 15 repeats, instead of the 12 repeats found in
CC   M129 strain.
CC   -!- MISCELLANEOUS: THE SEQUENCE SHOWN HERE IS THAT OF STRAIN FH.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AC013289; AAC52549.1; -;  
DR PIR; D96719; D96719.  
DR HSP; P15002; 184E.  
DR InterPro; IPR001731; Alad\_dehydratase.  
DR Pfam; PF00490; ALAD; 1.  
DR PRINTS; PR00144; DALDHYDRATASE.  
DR ProDom; PD002304; Alad\_dehydratase; 1.  
DR PROSITE; PS00169; D\_ALA\_DEHYDRATASE; 1.  
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Lyase; Magnesium;  
KW Chloroplast; Transit peptide.  
FT TRANSIT 1 ? CHLOROPLAST (BY SIMILARITY).  
FT CHAIN ? 430 DELTA-AMINOLEVULINIC ACID DEHYDRATASE.  
FT DOMAIN 219 237 MAGNESIUM-BINDING (BY SIMILARITY).  
FT ACT SITE 351 351 BY SIMILARITY.  
SQ SEQUENCE 430 AA; 46690 MW; 44B0984247FC6147 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 430;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LLRDKSP 267  
|||||  
DB 362 LLRDKSP 368

RESULT 30  
GLNA\_RHOCA  
ID \_GLNA\_RHOCA STANDARD; PRT; 438 AA.  
AC P13499;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase)  
DE (Fragment).  
GN GLNA.  
OS Rhodospirillum rubrum (Rhodospirillum rubrum capsulata).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;  
OC Rhodobacteraceae; Rhodobacter.  
OX NCBI\_TaxID=1061;  
RN [1]  
RP SEQUENCE OF 1-130 FROM N.A.  
RX MEDLINE=90094270; PubMed=2152916;  
RA Kranz R.G., Pace V.M., Caldwell I.M.;  
RT "Inactivation, sequence, and lacZ fusion analysis of a regulatory  
RT locus required for repression of nitrogen fixation genes in  
RT Rhodobacter capsulatus";  
RL J. Bacteriol. 172:53-62 (1990).  
RN [2]  
RP SEQUENCE OF 31-438 FROM N.A.  
RC STRAIN=BI00;  
RX MEDLINE=95362690; PubMed=7635841;  
RA Borghese R., Wall J.D.;  
RT "Regulation of the glnBA operon of Rhodobacter capsulatus";  
RL J. Bacteriol. 177:4549-4552 (1995).  
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +  
CC L-glutamine.  
CC -!- ENZYME REGULATION: The activity of this enzyme is controlled by  
CC adenylation. The fully adenylation enzyme complex is inactive.  
CC -!- SUBUNIT: Oligomer of 12 subunits arranged in the form of two  
CC hexagons.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the glutamine synthetase family.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----

DR EMBL; U25953; AAA87025.1; -;  
DR HSP; M28244; AAA26123.1; -;  
DR InterPro; IPR008147; Gln\_synt\_beta.  
DR InterPro; IPR008146; Gln\_synt\_C.  
DR InterPro; IPR004809; GlnA.  
DR Pfam; PF00120; gln-synt; 1.  
DR Pfam; PF03951; gln-synt\_N; 1.  
DR ProDom; PD001057; Gln\_synt\_C; 1.  
DR TIGRFAMs; TIGR00653; GlnA; 1.  
DR PROSITE; PS00180; GLNA 1; 1.  
DR PROSITE; PS00181; GLNA\_ATP; 1.  
KW Nitrogen fixation; Ligase.  
FT BINDING 397 397 AMP (UNDER CONDITIONS OF ABUNDANT  
FT GLUTAMINE) (BY SIMILARITY).  
FT NON TER 438 438  
SQ SEQUENCE 438 AA; 48391 MW; 07B81EBF6534991F CRC64;

Query Match 2.1%; Score 7; DB 1; Length 438;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 VTLLADL 86  
|||||  
DB 33 VTLLADL 39

RESULT 31  
YB8G\_YEAST  
ID \_YB8G\_YEAST STANDARD; PRT; 474 AA.  
AC P38358;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 51.7 kDa protein in Ctrp1-SUL2 intergenic region.  
GN YBR293W OR YBR2109.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RA Fritz C., Hollenberg C.P., Kirchner L., Rad M.R.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN  
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; Z36162; CAA85258.1; -;  
DR PIR; S46175; S46175.  
DR GerMOnline; 138836; -;  
DR SGD; S0000497; YBR293W.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR Pfam; PF00083; sugar\_tr; 1.  
DR PROSITE; PS50850; MFS; 1.  
KW Hypothetical protein; Transport; Transmembrane.  
FT TRANSMEM 34 54 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 98 118 POTENTIAL.  
FT TRANSMEM 122 142 POTENTIAL.  
FT TRANSMEM 168 188 POTENTIAL.

```

FT TRANSMEM 197 217 POTENTIAL.
FT TRANSMEM 239 259 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 376 396 POTENTIAL.
FT TRANSMEM 448 468 POTENTIAL.
SQ SEQUENCE 474 AA; 51677 MW; 1F3FDC877795ACFA CRC64;

Query Match 2.1%; Score 7; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 ILKDNLA 26
Db 462 ILKDNLA 468

RESULT 32
MYOC MOUSE
ID MYOC MOUSE STANDARD; PRT; 490 AA.
AC 070624; 070289;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myocilin precursor (Trabecular meshwork-induced glucocorticoid
DE response protein).
GN MYOC OR TIGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=129/Sv;
RX MEDLINE=98345432; PubMed=9680392;
RA Abderrahim H., Jaramillo-Babb V.L., Zhou Z., Vollrath D.;
RT "Characterization of the murine TIGR/myocilin gene.";
RL Mamm. Genome 9:673-675(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98217378; PubMed=9548973;
RA Fingert J.H., Ying L., Swiderski R.E., Nystuen A.M., Arbour N.C.,
RA Alward W.L.M., Sheffield V.C., Stone E.M.;
RT "Characterization and comparison of the human and mouse GLCIA glaucoma
RT genes.";
RL Genome Res. 8:377-384(1998).
RN [3]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND VARIANT ALA-164.
RC STRAIN=BALB/cJ, C3H/HeJ, and C57BL/6J; TISSUE=Brain, and Muscle;
RX MEDLINE=98249809; PubMed=9588210;
RA Tomarev S.I., Tamm E.R., Chang B.;
RT "Characterization of the mouse Myoc/Tigr gene.";
RL Biochem. Biophys. Res. Commun. 245:887-893(1998).
RN [4]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=BALB/c; TISSUE=Skeletal muscle;
RX MEDLINE=98340858; PubMed=9675094;
RA Takahashi H., Noda S., Imamura Y., Nagasawa A., Kubota R., Mashima Y.,
RA Kudoh J., Oguchi Y., Shimizu N.;
RT "Mouse myocilin (Myoc) gene expression in ocular tissues.";
RL Biochem. Biophys. Res. Commun. 248:104-109(1998).
RN [5]
RP INTERACTION WITH OLFM3.
RX MEDLINE=22013900; PubMed=12019210;
RA Torrado M., Trivedi R., Zinovieva R., Karavanova I., Tomarev S.I.;
RT "Optimedin; a novel olfactomedin-related protein that interacts with
RT myocilin.";
RL Hum. Mol. Genet. 11:1291-1301(2002).
CC -!- FUNCTION: May participate in the obstruction of fluid outflow in
CC the trabecular meshwork.
CC -!- SUBUNIT: Homodimer. Interacts with OLFM3.
CC -!- SUBCELLULAR LOCATION: Located preferentially in the ciliary

```

```

rootlet and basal body of the connecting cilium of photoreceptor
cells, and in the rough endoplasmic reticulum. Also secreted.
-!- TISSUE SPECIFICITY: Expressed in ciliary body, iris, retina,
trabecular network and sclera but not in lens or cornea. Also
expressed strongly in skeletal muscle and weakly in heart, brain,
testis, liver, kidney, thyroid and epididymis. No expression
detected in embryo.
-!- POLYMORPHISM: Variant Ala-164 is found in strain BALB/cJ which has
a low intraocular pressure. Variant Thr-164 is found in strains
C3H/HeJ and C57BL/6J, two strains which have a relatively high
intraocular pressure.
-!- SIMILARITY: Belongs to the olfactomedin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AF041335; AAC32805.1; -
EMBL; AF041333; AAC32805.1; JOINED.
EMBL; AF041334; AAC32805.1; JOINED.
EMBL; AF049796; AAC14265.1; -
EMBL; AF049795; AAC14265.1; JOINED.
EMBL; AF049794; AAC14265.1; JOINED.
EMBL; AF039869; AAC40112.1; -
EMBL; AB013592; BAA32031.1; -
PIR; JE0096; JEQ096.
MGD; MGI:1202864; Myoc.
GO; GO:0005615; C:extracellular space; IDA.
GO; GO:0005515; F:protein binding; IPI.
InterPro; IPR003112; Olfac_like.
PIfam; PF02191; OLF; 1.
SMART; SM00284; OLF; 1.
Coiled coil; Glycoprotein; Signal; Polymorphism.
SIGNAL 1 18 POTENTIAL.
CHAIN 19 490 MYOCILIN.
DOMAIN 69 170 COILED COIL (POTENTIAL).
DOMAIN 231 490 OLFACTOMEDIN-LIKE.
DISULFID 231 419 BY SIMILARITY.
CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
VARIANT 164 164 T -> A (IN STRAIN BALB/CJ).
SQ SEQUENCE 490 AA; 55314 MW; 2F090571E97B0425 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 260 NLLRDKS 266
Db 130 NLLRDKS 136
|||||
|||||

RESULT 33
CF26_BRARE
ID CP26_BRARE STANDARD; PRT; 492 AA.
AC P79739;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 26 (EC 1.14.-.-) (Retinoic acid-metabolizing
DE cytochrome) (P450RA1) (Retinoic acid 4-hydroxylase).
DE Cytochrome P450 26.
GN CYP26A1 OR CYP26.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97094702; PubMed=8939936;

```

[3]  
SEQUENCE FROM N.A.  
STRAIN=Berkeley;  
MEDLINE=20196066; PubMed=10731132;  
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
R.A. Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Ran K.H., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
Abrill J.F., Agbayani A., An H.-U., Andrews-Frannkoch C., Baldwin D.,  
Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,  
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
Lasko P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,  
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.;  
"The genome sequence of *Drosophila melanogaster*."  
Science 287:2185-2195(2000).  
[4]  
SEQUENCE FROM N.A.  
STRAIN=Berkeley; TISSUE=Embryo;  
MEDLINE=22426066; PubMed=12537569;  
Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
Rubin G.M., Celniker S.E.;  
"A *Drosophila* full-length cDNA resource."  
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
[5]  
FUNCTION.  
MEDLINE=21850185; PubMed=11861468;  
Micchelli C.A., The I., Selva E., Mogila V., Perrimon N.;  
"Rasp, a putative transmembrane acyltransferase, is required for  
Hedgehog signaling."  
Development 129:843-851(2002).  
-!- FUNCTION: Required in hedgehog (hh) expressing cells for  
production of appropriate signaling activity in embryos and in the  
imaginal precursors of adult tissues. Acts within the secretory  
pathway to catalyze amino-terminal palmitoylation of Hh; this  
lipid modification is required for the embryonic and larval  
patterning activities of the Hh signal. Not required for Wg  
signaling.  
-!- SUBCELLULAR LOCATION: Integral membrane protein.  
-!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.  
-!- SIMILARITY: Belongs to the membrane-bound acyltransferase family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
altered.

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; AF393157; AAK73748.1; -;  
 CC DR EMBL; AF398410; AAK97480.1; -;  
 CC DR EMBL; AE003477; AAF47725.1; -;  
 CC DR EMBL; AY119202; AAM51062.1; -;  
 CC DR FlyBase; FBgn0024134; rasp.  
 CC DR GO; GO:0007225; P:patched receptor ligand processing; IMP.  
 CC DR InterPro; IPR004299; MBOAT fam.  
 CC DR Pfam; PF03062; MBOAT; 1.  
 CC KW Transferrase; Acyltransferase; Developmental protein;  
 CC KW Segmentation polarity protein; Transmembrane.  
 CC FT TRANSMEM 15 35 POTENTIAL.  
 CC FT TRANSMEM 73 93 POTENTIAL.  
 CC FT TRANSMEM 105 125 POTENTIAL.  
 CC FT TRANSMEM 134 154 POTENTIAL.  
 CC FT TRANSMEM 206 226 POTENTIAL.  
 CC FT TRANSMEM 243 263 POTENTIAL.  
 CC FT TRANSMEM 293 313 POTENTIAL.  
 CC FT TRANSMEM 372 392 POTENTIAL.  
 CC FT TRANSMEM 429 449 POTENTIAL.  
 CC FT TRANSMEM 461 481 POTENTIAL.  
 CC FT ACT SITE 381 381 POTENTIAL.  
 CC FT CONFLICT 91 91 G -> V (IN REF. 1 AND 4).  
 CC SQ SEQUENCE 500 AA; 58105 MW; 4498DC306976F2A2 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84  
 DB 124 LLVTLIA 130

RESULT 35  
 MYOC\_RAT  
 ID MYOC\_RAT STANDARD; PRT; 502 AA.  
 AC Q9RLJ4; Q9ZY4; 502 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Myocilin precursor (Trabecular meshwork-induced glucocorticoid  
 DE response protein).  
 GN MYOC OR TIGR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP STRAIN=Wistar Kyoto; TISSUE=Eye;  
 RC Yaeoda K., Yamamoto T., Funaki H., Koyama Y., Nihei K., Tani T.,  
 RA Yaeita E., Kawasaki K., Abe H., Kihara I.;  
 RT "Molecular cloning of myocilin gene from rat eye."  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.; AND TISSUE SPECIFICITY.  
 RC STRAIN=Sprague-Dawley; TISSUE=Eye;  
 RX MEDLINE=20295350; PubMed=1083334;  
 RA Taguchi M., Kanno H., Kubota R., Miwa S., Shishiba Y., Ozawa Y.;  
 RT "Molecular cloning and expression profile of rat myocilin."  
 RL Mol. Genet. Metab. 70:75-80(2000).  
 CC -!- FUNCTION: May participate in the obstruction of fluid outflow in  
 CC the trabecular meshwork (By similarity).  
 CC -!- SUBUNIT: Homodimer. Interacts with MYOC (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Located preferentially in the ciliary  
 CC rootlet and basal body of the connecting cilium of photoreceptor  
 CC cells, and in the rough endoplasmic reticulum. Also secreted (By  
 CC similarity).  
 CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and

CC retina. Also detected at lower levels in thyroid gland but not in  
 CC other endocrine glands such as the adrenal or pituitary  
 CC glands.  
 CC -!- SIMILARITY: Belongs to the olfactomedin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF093567; BAA46401.1; -;  
 CC DR EMBL; AB013993; BAA34199.1; -;  
 CC DR InterPro; IPR003112; Olfac\_like.  
 CC DR Pfam; PF02191; OLF; 1.  
 CC DR SMART; SM00284; OLF; 1.  
 CC KW Coiled coil; Glycoprotein; Signal.  
 CC FT SIGNAL 1 31 POTENTIAL.  
 CC FT CHAIN 32 502 MYOCILIN.  
 CC FT DOMAIN 82 183 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 243 502 OLFACTOMEDIN-LIKE.  
 CC FT DISULFID 243 431 BY SIMILARITY.  
 CC FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CONFLICT 6 6 Y -> R (IN REF. 2).  
 CC FT CONFLICT 329 329 S -> A (IN REF. 2).  
 CC SQ SEQUENCE 502 AA; 56442 MW; 2FE8FBE53CF48BBA CRC64;

Query Match 2.1%; Score 7; DB 1; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 DB 143 NLLRDKS 149

RESULT 36  
 MYOC\_HUMAN  
 ID MYOC\_HUMAN STANDARD; PRT; 504 AA.  
 AC Q99972; O00620;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Myocilin precursor (Trabecular meshwork-induced glucocorticoid  
 DE response protein).  
 GN MYOC OR TIGR OR GLC1A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.; AND VARIANTS JOAG VAL-364 AND HIS-437.  
 RX MEDLINE=97158493; PubMed=9005853;  
 RA Stone E.M., Fingert J.H., Alward W.L.M., Nguyen T.D., Polansky J.R.,  
 RA Sunden S.L.F., Nishimura D., Clark A.F., Nyström A., Nichols B.E.,  
 RA Mackey D.A., Ritch R., Kalenak J.W., Craven E.R., Sheffield V.C.;  
 RT "Identification of a gene that causes primary open angle glaucoma."  
 RL Science 275:668-670(1997).  
 RN [2]  
 RP REVISIONS, AND SEQUENCE OF 1-6 AND 33-37.  
 RX MEDLINE=98165818; PubMed=9497363;  
 RA Nguyen T.D., Chen P., Huang W.D., Chen H., Johnson D., Polansky J.R.;  
 RT "gene structure and properties of TIGR, an olfactomedin-related  
 RT glycoprotein cloned from glucocorticoid-induced trabecular meshwork  
 RT cells."  
 RL J. Biol. Chem. 273:6341-6350(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97424389; PubMed=9280311;  
 RA Ortego J., Escibano J., Coca-Prados M.;  
 RT "Cloning and characterization of subtracted cDNAs from a human



RT ciliary body library encoding TIGR, a protein involved in juvenile  
 RT open angle glaucoma with homology to myosin and olfactomedin.";  
 RL FEBS Lett. 413:349-353(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=97312692; PubMed=9169133;  
 RA Kubota R., Noda S., Wang Y., Minoshima S., Asakawa S., Kudoh J.,  
 RA Mashima Y., Oguchi Y., Shimizu N.;  
 RT "A novel myosin-like protein (myocilin) expressed in the connecting  
 RT cilium of the photoreceptor: molecular cloning, tissue expression,  
 RT and chromosomal mapping.";  
 RL Genomics 41:360-369(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX Garchon H.-J.;  
 RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RL [6]  
 RP SEQUENCE FROM N.A.  
 RA Deadman R.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98217378; PubMed=9548973;  
 RA Fingert J.H., Ying L., Swiderski R.E., Nystuen A.M., Arbour N.C.,  
 RA Alward W.L.M., Sheffield V.C., Stone E.M.;  
 RT "Characterization and comparison of the human and mouse GLC1A glaucoma  
 RT genes.";  
 RL Genome Res. 8:377-384(1998).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98113364; PubMed=9446806;  
 RA Kubota R., Kudoh J., Mashima Y., Asakawa S., Minoshima S.,  
 RA Fujitani K., Oguchi Y., Shimizu N.;  
 RT "Genomic organization of the human myocilin gene (MYOC) responsible  
 RT for primary open angle glaucoma (GLC1A).";  
 RL Biochem. Biophys. Res. Commun. 242:396-400(1998).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=92388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [10]  
 RP OLFACTOMEDIN DOMAIN DISULFIDE BOND.  
 RX MEDLINE=22501919; PubMed=12615070;  
 RA Nagy I., Trexler M., Pothly L.;  
 RT "Expression and characterization of the olfactomedin domain of human  
 RT myocilin.";  
 RL Biochem. Biophys. Res. Commun. 302:554-561(2003).  
 RN [11]  
 RP INTERACTION WITH OLFM3.  
 RX MEDLINE=22013900; PubMed=12019210;  
 RA Torrado M., Trivedi R., Zinovieva R., Karavanova I., Tomarev S.I.;

"Optimedin: a novel olfactomedin-related protein that interacts with  
 myocilin.";  
 Hum. Mol. Genet. 11:1291-1301(2002).  
 RN [12]  
 RP VARIANTS JOAG ARG-367 AND LEU-370.  
 RX MEDLINE=9808006; PubMed=9345106;  
 RA Suzuki Y., Shirato S., Taniguchi F., Ohara K., Nishimaki K., Ohta S.;  
 RT "Mutations in the TIGR gene in familial primary open-angle glaucoma  
 RT in Japan.";  
 RL Am. J. Hum. Genet. 61:1202-1204(1997).  
 RN [13]  
 RP VARIANTS POAG ARG-246; LEU-370; SER-477; LYS-480 AND PHE-499.  
 RX MEDLINE=97472461; PubMed=9328473;  
 RA Adam M.F., Belmondou A., Binieta P., Brezin A.P., Valtot F.,  
 RA Bechetolle A., Descotte J.-C., Copin B., Gomez L., Chaventre A.,  
 RA Bach J.-F., Garchon H.-J.;  
 RT "Recurrent mutations in a single exon encoding the evolutionarily  
 RT conserved olfactomedin-homology domain of TIGR in familial open-angle  
 RT glaucoma.";  
 RL Hum. Mol. Genet. 6:2091-2097(1997).  
 RN [14]  
 RP VARIANTS JOAG ARG-337.  
 RX MEDLINE=98027214; PubMed=9361308;  
 RA Stoilova D., Child A., Brice G., Crick R.P., Fleck B.W., Sarfarazi M.;  
 RT "Identification of a new 'TIGR' mutation in a family with juvenile-  
 RT onset primary open angle glaucoma.";  
 RL Ophthalmic Genet. 18:109-118(1997).  
 RN [15]  
 RP VARIANTS COAG LYS-352 AND MET-377, AND VARIANTS JOAG LEU-370 AND  
 RP HIS-437.  
 RX MEDLINE=99011274; PubMed=9792882;  
 RA Wiggs J.L., Allingham R.R., Vollrath D., Jones K.H., De La Paz M.,  
 RA Kern J., Patterson K., Babb V.L., Del Bono E.A., Broomer B.W.,  
 RA Pericak-Vance M.A., Haines J.L.;  
 RT "Prevalence of mutations in TIGR/Myocilin in patients with adult and  
 RT juvenile primary open-angle glaucoma.";  
 RL Am. J. Hum. Genet. 63:1549-1552(1998).  
 RN [16]  
 RP VARIANTS JOAG ARG-367 AND LEU-370.  
 RX MEDLINE=98141135; PubMed=9490287;  
 RA Polansky J., Nguyen T., Timmerman C., Mardin C.Y., Budde W.M., Liehr T.,  
 RA Polansky J., Nguyen T., Timmerman C., Mardin C.Y., van Broeckhoven C.,  
 RA Naumann G.O.H., Pfeiffer R.A., Rautenstrauss B.W.;  
 RT "Juvenile open angle glaucoma: fine mapping of the TIGR gene to  
 RT 1q24.3-q25.2 and mutation analysis.";  
 RL Hum. Genet. 102:103-106(1998).  
 RN [17]  
 RP VARIANTS COAG ARG-367, AND VARIANT JOAG PHE-426.  
 RX MEDLINE=98180724; PubMed=9521427;  
 RA Mansergh F.C., Kenna P.F., Ayuso C., Kiang A.-S., Humphries P.,  
 RA Farrar G.J.;  
 RT "Novel mutations in the TIGR gene in early and late onset open angle  
 RT glaucoma.";  
 RL Hum. Mutat. 11:244-251(1998).  
 RN [18]  
 RP VARIANTS JOAG LEU-370; ALA-380 AND PRO-502, AND VARIANT LYS-76.  
 RX MEDLINE=99079298; PubMed=9863594;  
 RA Stoilova D., Child A., Brice G., Desai T., Barsom-Homsy M.,  
 RA Ozdemir N., Chevrete L., Adam M.F., Garchon H.-J., Pitts Crick R.,  
 RA Sarfarazi M.;  
 RT "Novel TIGR/MYOC mutations in families with juvenile onset primary  
 RT open angle glaucoma.";  
 RL J. Med. Genet. 35:989-992(1998).  
 RN [19]  
 RP VARIANTS POAG GLU-423.  
 RX MEDLINE=98361153; PubMed=9697688;  
 RA Morissette J., Clepet C., Moisan S., Dubois S., Winstall E.,  
 RA Vermeeren D., Nguyen T.D., Polansky J.R., Cote G., Anctil J.-L.,  
 RA Anyot M., Plante M., Falardeau P., Raymond V.;  
 RT "Homozygotes carrying an autosomal dominant TIGR mutation do not  
 RT manifest glaucoma.";  
 RL Nat. Genet. 19:319-321(1998).  
 RN [20]



RESULT 38



FT TRANSNM 736 756 POTENTIAL.  
 FT TRANSNM 768 788 POTENTIAL.  
 FT TRANSNM 803 823 POTENTIAL.  
 FT TRANSNM 850 870 POTENTIAL.  
 FT TRANSNM 899 919 POTENTIAL.  
 FT TRANSNM 162 162 (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 872 872 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 923 AA; 102549 MW; D168FC52FE26C5CB CRC64;

Query Match 2.1%; Score 7; DB 1; Length 923;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 75 SSSLVLT 81  
 |||||  
 Db 784 SSSLVLT 790

RESULT 40  
 NME3 RAT  
 ID NME3 RAT STANDARD; PRT; 1237 AA.  
 AC Q00961;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glutamate [NMDA] receptor subunit epsilon 3 precursor (N-methyl  
 D-aspartate receptor subtype 2C) (NR2C) (NMDAR2C).  
 GN GRIN2C.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=92271257; PubMed=1350383;  
 RA Monyer H., Sprengel R., Schoepfer R., Herb A., Higuchi M., Lomeli H.,  
 Burnashev N., Sakmann B., Seeburg P.H.;  
 RT "Heteromeric NMDA receptors: molecular and functional distinction of  
 subtypes.";  
 RL Science 256:1217-1221(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93155102; PubMed=8428958;  
 RA Ishii T., Moriyoshi K., Sugihara H., Sakurada K., Kadotani H.,  
 RA Yokoi M., Akazawa C., Shigemoto R., Mizuno N., Masu M.,  
 RA Nakanishi S.;  
 RT "Molecular characterization of the family of the N-methyl-D-aspartate  
 receptor subunits.";  
 RL J. Biol. Chem. 268:2836-2843(1993).  
 CC -!- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS  
 POSSESSES HIGH CALCIUM PERMEABILITY AND VOLTAGE-DEPENDENT  
 SENSITIVITY TO MAGNESIUM AND IS MODULATED BY GLYCINE.  
 CC -!- SUBUNIT: Heterodimer of an epsilon subunit and a zeta subunit.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M91563; AAA41713.1;  
 CC EMBL; D13212; EAA02499.1; ALT\_INIT.  
 CC HSP; P19491; IGR2.  
 CC InterPro; IPR001320; Ion\_glu\_receptor.  
 CC InterPro; IPR001508; NMDA\_receptor.  
 CC InterPro; IPR001311; SBP/glu\_receptor.

DR Pfam; PF00060; lig\_chan; 1.  
 DR PRINTS; PR00177; NMDARECEPTOR.  
 DR SMART; SM00079; PBPe; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane; Calcium; Magnesium.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 1237 GLUTAMATE [NMDA] RECEPTOR SUBUNIT  
 EPSILON 3.  
 FT DOMAIN 20 554 EXTRACELLULAR (POTENTIAL).  
 FT TRANSNM 555 574 1 (POTENTIAL).  
 FT TRANSNM 597 617 2 (POTENTIAL).  
 FT TRANSNM 627 652 3 (POTENTIAL).  
 FT TRANSNM 815 835 4 (POTENTIAL).  
 FT SITE 612 612 FUNCTIONAL DETERMINANT OF NMDA  
 RECEPTORS (BY SIMILARITY).  
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1237 AA; 135271 MW; B175993804B337A4 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 1237;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 63 TEAVAQL 69  
 |||||  
 Db 101 TEAVAQL 107

RESULT 41  
 NME3 MOUSE  
 ID NME3 MOUSE STANDARD; PRT; 1239 AA.  
 AC Q01038;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glutamate [NMDA] receptor subunit epsilon 3 precursor (N-methyl  
 D-aspartate receptor subtype 2C) (NR2C) (NMDAR2C).  
 GN GRIN2C.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93210564; PubMed=1377365;  
 RA Kutsuwada T., Kasaiwabuchi N., Mori H., Sakimura K., Kuehniya E.,  
 RA Araki K., Meguro H., Masaki H., Kumanishi T., Arakawa M.,  
 RA Mishina M.;  
 RT "Molecular diversity of the NMDA receptor channel.";  
 RL Nature 358:36-41(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93210564; PubMed=1377365;  
 RA Kutsuwada T., Kasaiwabuchi N., Mori H., Sakimura K., Kuehniya E.,  
 RA Araki K., Meguro H., Masaki H., Kumanishi T., Arakawa M.,  
 RA Mishina M.;  
 RT "Molecular diversity of the NMDA receptor channel.";  
 RL Nature 358:36-41(1992).  
 CC REVISIONS.  
 CC Kashiwabuchi N.;  
 CC Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: NMDA receptor subtype of glutamate-gated ion channels  
 possesses high calcium permeability and voltage-dependent  
 sensitivity to magnesium and is mediated by glycine.  
 CC -!- SUBUNIT: Heterodimer of an epsilon subunit and a zeta subunit.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D10694; BAA01536.1;  
 CC FIR; I49705; I49705.

```

DR HSP; P19491; 1GR2;
DR MGD; MGI:95822; Grin2c;
DR InterPro; IPR001320; Ion_glu_receptor;
DR InterPro; IPR001508; NMDA_receptor;
DR InterPro; IPR001311; SBP/Glu_receptor;
DR Pfam; PF00060; lig_chan; 1;
DR PRINTS; PR00177; NMDARECEPTOR;
DR SMART; SM00079; PBPe; 1;
KW Receptor; Signal; Transmembrane; Postsynaptic membrane; Calcium;
KW Ionic channel; Magnesium;
FT SIGNAL 1 19
FT CHAIN 20 1239
FT DOMAIN 20 553
FT TRANSMEM 554 574
FT TRANSMEM 597 617
FT TRANSMEM 627 647
FT TRANSMEM 815 835
FT SITE 612 612
FT CARBOHYD 70 70
FT CARBOHYD 337 337
FT CARBOHYD 438 438
FT CARBOHYD 539 539
SQ SEQUENCE 1239 AA; 135420 MW; 793B8E731E20C3C9 CRC64;

Query Match 2.1%; Score 7; DB 1; Length 1239;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEVAQL 69
DB 101 TEVAQL 107

RESULT 42
RNG2 SCHPO STANDARD; PRT; 1489 AA.
AC 014188; Q9USG0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras GTPase-activating-like protein rng2 (Ring assembly protein 2).
GN RNG2 OR SPAC4F8.13C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Ruter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

```

```

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe I., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 354-531 FROM N.A., AND SUBCELLULAR LOCATION.
RC STRAIN=968 h90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
RN [3]
RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RC MEDLINE=98298927; PubMed=9635188;
RA Eng K., Naqvi N.I., Wong K.C.Y., Balasubramanian M.K.;
RT "Rng2p, a protein required for cytokinesis in fission yeast, is a
RT component of the actomyosin ring and the spindle pole body.";
RL Curr. Biol. 8:611-621(1998).
CC -!- FUNCTION: Required for cytokinesis. Component of the contractile
CC F-actin ring; required for its construction following assembly of
CC F-actin at the division site.
CC -!- SUBUNIT: Interacts with calmodulin.
CC -!- SUBCELLULAR LOCATION: Localized to the F-actin ring and spindle
CC pole body during interphase and mitosis. Also found in
CC septum.
CC -!- SIMILARITY: Contains 1 calponin-homology (CH) domain.
CC -!- SIMILARITY: Contains 6 IQ domains.
CC -!- SIMILARITY: Contains 1 Ras-GAP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z98530; CAB11059.1; -.
CC EMBL; AB027779; BAA87083.1; -.
CC PIR; T38842; T38842.
CC GeneDB_Spombe; SPAC4F8.13C; -.
CC GO; GO:0030480; C:contractile ring (sensu Fungi); IDA.
CC GO; GO:0030428; C:septum; IDA.
CC GO; GO:0005816; C:spindle pole body; IDA.
CC GO; GO:0005516; F:calmodulin binding; IPI.
CC InterPro; IPR001715; Calponin-like.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001936; RasGAP.
CC InterPro; IPR000593; RasGAP_C.
CC Pfam; PF00307; CH; 1.
CC Pfam; PF00612; IQ; 9.
CC Pfam; PF00616; RasGAP; 1.
CC Pfam; PF03836; RasGAP_C; 1.
CC ProDom; PD008735; RasGAP_C; 1.
CC SMART; SM00033; CH; 1.
CC SMART; SM00015; IQ; 4.
CC SMART; SM00323; RasGAP; 1.
CC PROSITE; PS50021; CH; 1.
CC PROSITE; PS50096; IQ; 6.
CC PROSITE; PS50018; Ras_GTPASE_ACTIV_2; 1.
KW Cell division; Septation; Calmodulin-binding; Repeat; Coiled coil.
FT DOMAIN 41 147
FT DOMAIN 359 388
FT DOMAIN 389 418
FT DOMAIN 418 449
FT DOMAIN 535 564
FT DOMAIN 565 594
FT DOMAIN 655 684

```

```
FT DOMAIN 734 770 COILED COIL (POTENTIAL).
FT DOMAIN 854 1077 RAS-GAP.
FT DOMAIN 1330 1364 COILED COIL (POTENTIAL).
SQ SEQUENCE 1489 AA; 171676 MW; 7D357FF9A7FED5EC CRC64;

Query Match 2.1%; Score 7; DB 1; Length 1489;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 HKVLVAD 200
Db 925 HKVLVAD 931

RESULT 43
RBP1 PLAVB STANDARD; PRT; 2869 AA.
AC Q00758;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulocyte binding protein 1 precursor.
GN RBP1.
OS Plasmodium vivax (strain Belm).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1] SEQUENCE FROM N.A.
RP MEDLINE=92315338; PubMed=1617731;
RX Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RA "A reticulocyte-binding protein complex of Plasmodium vivax
RT merzotiles."
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
CC human reticulocyte cells.
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M88097; AAA29743.1; -
KW Malaria; Receptor; Signal; Transmembrane.
FT SIGNAL 1 17
FT CHAIN 18 2869 RETICULOCYTE BINDING PROTEIN 1.
FT DOMAIN 18 2807 EXTRACELLULAR.
FT TRANSMEM 2808 2826 POTENTIAL.
FT DOMAIN 2827 2869 CYTOPLASMIC.
FT SITE 1030 1032 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 2869 AA; 330213 MW; B9DBE442205BCFF CRC64;

Query Match 2.1%; Score 7; DB 1; Length 2869;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLIGE 234
Db 404 SLKLIGE 410

RESULT 44
UTRO HUMAN STANDARD; PRT; 3433 AA.
ID UTRO HUMAN
AC F46939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE Utrophin (Dystrophin-related protein 1) (DRP1) (DRP).
GN UTREN OR DMDL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93096045; PubMed=1461283;
RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,
RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,
RA Edwards Y.H., Davies K.E.;
RA "Primary structure of dystrophin-related protein.";
RL Nature 360:591-593(1992).
RN [2]
RP INTERACTION WITH SNTB1.
RX MEDLINE=95146543; PubMed=7844150;
RA Ahn A.H., Kunkel L.M.;
RA "Syntrophin binds to an alternatively spliced exon of dystrophin.";
RL J. Cell Biol. 128:363-371(1995).
RN [3]
RP INTERACTION WITH SNTA1 AND SNTB2.
RX MEDLINE=96162017; PubMed=8576247;
RA Ahn A.H., Feener C.A., Gussoni E., Yoshida M., Ozawa E., Kunkel L.M.;
RA "The three human syntrophin genes are expressed in diverse tissues,
RA have distinct chromosomal locations, and each bind to dystrophin and
RA its relatives.";
RL J. Biol. Chem. 271:2724-2730(1996).
RN [4]
RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.
RP MEDLINE=99141377; PubMed=9887274;
RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,
RA Kendrick-Jones J.;
RA "Crystal structure of the actin-binding region of utrophin reveals a
RA head-to-tail dimer.";
RL Structure 7:1539-1546(1999).
CC -!- FUNCTION: May play a role in anchoring the cytoskeleton to the
CC plasma membrane (By similarity to dystrophin).
CC -!- SUBUNIT: Interacts with the syntrophins SNTA1; SNTB1 and SNTB2.
CC -!- SUBCELLULAR LOCATION: Neuromuscular junction.
CC -!- TISSUE SPECIFICITY: Muscle.
CC -!- SIMILARITY: STRONG, TO DYSTROPHIN.
CC -!- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
CC ABP-120, ABP-180, OR BETA-PODRIN).
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 1 WW domain.
CC -!- SIMILARITY: Contains 20 spectrin repeats.
CC -!- SIMILARITY: Contains 1 ZZ-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69086; CAA48829.1; -
DR PIR; S28381; S28381.
DR PDB; 1BHD; 16-FEB-99.
DR PDB; 1QAG; 01-JAN-00.
DR Genew; HGNC:12635; UTRN.
DR MIM; 128240; -.
DR GO; GO:0005856; C:cytoskeleton; TAS.
```

DR GO: GO:0005624; C:membrane fraction; TAS.  
 DR GO: GO:0005886; C:plasma membrane; TAS.  
 DR GO: GO:0006936; P:muscle contraction; TAS.  
 DR GO: GO:0007517; P:muscle development; TAS.  
 DR InterPro: IPR001589; Actbind.actinin.  
 DR InterPro: IPR001715; Calponin-like.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR001202; WW Rsp5.WWP.  
 DR InterPro: IPR000433; ZnF\_ZZ.  
 DR Pfam: PF00307; CH; 2.  
 DR Pfam: PF00435; spectrin; 18.  
 DR Pfam: PF00397; WW; 1.  
 DR Pfam: PF00569; ZZ; 1.  
 DR SMART: SM00033; CH; 2.  
 DR SMART: SM00150; SPEC; 19.  
 DR SMART: SM00456; WW; 1.  
 DR SMART: SM00291; ZnF\_ZZ; 1.  
 DR PROSITE: PS00019; ACTININ\_1; 1.  
 DR PROSITE: PS00020; ACTININ\_2; 1.  
 DR PROSITE: PS00021; CH; 2.  
 DR PROSITE: PS01159; WW DOMAIN 1; 1.  
 DR PROSITE: PS00020; WW DOMAIN 2; 1.  
 DR PROSITE: PS01357; ZE\_ZZ\_1; 1.  
 DR PROSITE: PS01357; ZE\_ZZ\_2; 1.  
 DR PROSITE: PS01357; ZE\_ZZ\_3; 1.  
 DR Structural protein; Actin-binding; Cytoskeleton;  
 KW Repeat; 3D-structure; Zinc-finger.  
 FT DOMAIN 1 246 ACTIN-BINDING.  
 FT DOMAIN 31 135 CH 1.  
 FT DOMAIN 150 252 CH 2.  
 FT REPEAT 253 308 SPECTRIN 1.  
 FT REPEAT 309 417 SPECTRIN 2.  
 FT REPEAT 418 526 SPECTRIN 3.  
 FT REPEAT 541 637 SPECTRIN 4.  
 FT REPEAT 687 798 SPECTRIN 5.  
 FT REPEAT 803 902 SPECTRIN 6.  
 FT REPEAT 1016 1083 SPECTRIN 7.  
 FT REPEAT 1125 1230 SPECTRIN 8.  
 FT REPEAT 1248 1334 SPECTRIN 9.  
 FT REPEAT 1432 1541 SPECTRIN 10.  
 FT REPEAT 1544 1649 SPECTRIN 11.  
 FT REPEAT 1652 1753 SPECTRIN 12.  
 FT REPEAT 1910 1968 SPECTRIN 13.  
 FT REPEAT 1976 2081 SPECTRIN 14.  
 FT REPEAT 2258 2333 SPECTRIN 15.  
 FT REPEAT 2399 2440 SPECTRIN 16.  
 FT REPEAT 2443 2556 SPECTRIN 17.  
 FT REPEAT 2559 2636 SPECTRIN 18.  
 FT REPEAT 2658 2888 SPECTRIN 19.  
 FT REPEAT 2691 2797 SPECTRIN 20.  
 FT DOMAIN 2812 2845 WW.  
 FT ZN\_FING 3064 3111 ZZ-TYPE.  
 FT HELIX 32 46  
 FT TURN 47 49  
 FT TURN 56 62  
 FT TURN 64 74  
 FT TURN 75 75  
 FT HELIX 86 102  
 FT TURN 103 104  
 FT TURN 112 117  
 FT TURN 118 118  
 FT HELIX 120 134  
 FT TURN 135 135  
 FT HELIX 136 149  
 FT TURN 152 163  
 FT TURN 164 164  
 FT TURN 168 169  
 FT TURN 177 179  
 FT TURN 180 181  
 FT HELIX 183 191  
 FT TURN 192 192  
 FT TURN 194 196  
 FT HELIX 199 204  
 FT HELIX 207 222

FT HELIX 230 233  
 FT HELIX 240 252  
 FT TURN 253 253  
 SQ SEQUENCE 3433 AA; 394488 MW; EAE8DB409F858E5B CRC64;  
 Query Match 2.1%; Score 7; DB 1; Length 3433;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 VKILKDN 24  
 Db 1182 VKILKDN 1188  
 RESULT 45  
 LYST\_MOUSE STANDARD; PRT; 3788 AA.  
 ID AC P97412; Q62403; Q8VBSE6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Lysosomal trafficking regulator (Beige protein) (CHS1 homolog).  
 GN CHS1 OR LYST OR BG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=97358584; PubMed=9215680;  
 RA Barbosa M.D.F.S., Barrat F.J., Tchernev V.T., Nguyen Q.A.,  
 RA Mishra V.S., Colman S.D., Pastural E., Dufourcq-Iagelouse R.,  
 RA Fischer A., Holcombe R.F., Wallace M.R., Brandt S.J.,  
 RA De Saint Basile G., Kingmore S.F.;  
 RT "Identification of mutations in two major mRNA isoforms of the  
 RT Chediak-Higashi syndrome gene in human and mouse.";  
 RL Hum. Mol. Genet. 6:1091-1098(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX STRAIN=C57BL/6J;  
 RA MEDLINE=96353977; PubMed=8717042;  
 RA Barbosa M.D.F.S., Nguyen Q.A., Tchernev V.T., Ashley J.A.,  
 RA Dettler J.C., Blaydes S.M., Brandt S.J., Chotai D., Hodgman C.,  
 RA Solari R.C.E.S., Lovett M., Kingmore S.F.;  
 RT "Identification of the homologous beige and Chediak-Higashi syndrome  
 RT genes.";  
 RL Nature 382:262-265(1996).  
 RN [3]  
 RP ERRATUM.  
 RA Barbosa M.D.F.S., Nguyen Q.A., Tchernev V.T., Ashley J.A.,  
 RA Dettler J.C., Blaydes S.M., Brandt S.J., Chotai D., Hodgman C.,  
 RA Solari R.C.E.S., Lovett M., Kingmore S.F.;  
 RL Nature 385:97-97(1997).  
 RN [4]  
 RP SEQUENCE OF 1428-3788 FROM N.A. (ISOFORM 1).  
 RX STRAIN=C57BL/6J;  
 RA MEDLINE=96259558; PubMed=8673129;  
 RA Perou C.M., Moore K.J., Nagle D.L., Misumi D.J., Woolf E.A.,  
 RA McGrail S.H., Holmgren L., Brody T.B., Dussault B.J., Monroe C.A.,  
 RA Duyk G.M., Pryor R.J., Li L., Justice M.J., Kaplan J.;  
 RT "Identification of the murine beige gene by YAC complementation and  
 RT positional cloning.";  
 RL Nat. Genet. 13:303-308(1996).  
 CC -1- FUNCTION: May be required for sorting endosomal resident proteins  
 CC into late multivesicular endosomes by a mechanism involving  
 CC microtubules.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=P97412-1; Sequence=displayed;  
 CC Name=2;





```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RROIGT 111
Db 23 RROIGT 28

RESULT 48
BD07_BOVIN
ID BD07_BOVIN STANDARD; PRT; 40 AA.
AC P46165;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Beta-defensin 7 (BNDB-7) (BNDB-7).
GN DEFB7.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

SEQUENCE.
RP STRAIN=Hereford; TISSUE=Neutrophils;
RX MEDLINE=93203264; PubMed=8454635;
RA Salsted M.E., Tang Y.-Q., Morris W.L., McGuire P.A., Novotny M.J.,
RA Smith W., Henschen A.H., Cullor J.S.;
RT "Purification, primary structures, and antibacterial activities of
RT beta-defensins, a new family of antimicrobial peptides from bovine
RT neutrophils.";
RL J. Biol. Chem. 268:6641-6648 (1993).
CC -!- FUNCTION: Has bactericidal activity. Active against E.coli ML35
CC and S.aureus 502A.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neutrophilic granules.
CC -!- SIMILARITY: Belongs to the beta-defensin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF016394; AAC48801.1; -.
CC HSPSP; P46170; 1BNB.
CC InterPro; IPR001855; Defensin_beta.
CC Pfam; PF00711; Defensin_beta; 1.
CC SMART; SM00048; DEFSN; 1.
CC Antibioc; Signal; Pyrrolidone carboxylic acid.
CC NON_TER 1 1
CC SIGNAL <1 ? POTENTIAL.
CC PROPEP ? 15 POTENTIAL.
CC CHAIN 16 55 BETA-DEFENSIN 9.
CC MOD_RES 16 16 PYRROLIDONE CARBOXYLIC ACID.
CC DISULFID 24 53 BY SIMILARITY.
CC FT DISULFID 31 46 BY SIMILARITY.
CC FT DISULFID 36 54 BY SIMILARITY.
CC FT CONFLICT 48 48 A -> G (IN REF. 2).
CC SQ SEQUENCE 55 AA; 6049 MW; 48DAE6917DE366F2 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RROIGT 111
Db 25 RROIGT 30

RESULT 49
BD09_BOVIN
ID BD09_BOVIN STANDARD; PRT; 55 AA.
AC P46167; O18814;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-defensin 9 precursor (BNDB-9) (BNDB-9) (Fragment).
GN DEFB9 OR BNDB9.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
```

```
SEQUENCE FROM N.A.
TISSUE=Small intestine;
MEDLINE=98147718; PubMed=9488394;
Tarver A.P., Clark D.P., Diamond G., Russell J.P., Jones D.E., Sweeney R.W.,
Brdjument-Bromage H., Tempst P., Cohen K.S., Morris M.J.,
Wines M., Hwang S., Bevins C.L.;
"Enteric beta-defensin: molecular cloning and characterization of a
gene with inducible intestinal epithelial cell expression associated
with Cryptosporidium parvum infection.";
Infect. Immun. 66:1045-1056 (1998).
RN [2]

SEQUENCE OF 16-55.
RP STRAIN=Hereford; TISSUE=Neutrophils;
RX MEDLINE=93203264; PubMed=8454635;
RA Salsted M.E., Tang Y.-Q., Morris W.L., McGuire P.A., Novotny M.J.,
RA Smith W., Henschen A.H., Cullor J.S.;
RT "Purification, primary structures, and antibacterial activities of
RT beta-defensins, a new family of antimicrobial peptides from bovine
RT neutrophils.";
RL J. Biol. Chem. 268:6641-6648 (1993).
CC -!- FUNCTION: Has bactericidal activity. Active against E.coli ML35
CC and S.aureus 502A.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neutrophilic granules.
CC -!- SIMILARITY: Belongs to the beta-defensin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF016394; AAC48801.1; -.
CC HSPSP; P46170; 1BNB.
CC InterPro; IPR001855; Defensin_beta.
CC Pfam; PF00711; Defensin_beta; 1.
CC SMART; SM00048; DEFSN; 1.
CC Antibioc; Signal; Pyrrolidone carboxylic acid.
CC NON_TER 1 1
CC SIGNAL <1 ? POTENTIAL.
CC PROPEP ? 15 POTENTIAL.
CC CHAIN 16 55 BETA-DEFENSIN 9.
CC MOD_RES 16 16 PYRROLIDONE CARBOXYLIC ACID.
CC FT DISULFID 24 53 BY SIMILARITY.
CC FT DISULFID 31 46 BY SIMILARITY.
CC FT DISULFID 36 54 BY SIMILARITY.
CC FT CONFLICT 48 48 A -> G (IN REF. 2).
CC SQ SEQUENCE 55 AA; 6049 MW; 48DAE6917DE366F2 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RROIGT 111
Db 40 RROIGT 45

RESULT 50
RL29_TREPA
ID RL29_TREPA STANDARD; PRT; 72 AA.
AC O83227;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L29.
GN RPL29 OR RPL197.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
```

RN SEQUENCE FROM N.A.  
RC STRAIN=Nichols;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren E., Hardham J.W., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howland J.K., Chidambaram M., Utterback T.,  
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
RT spirochete.";  
RL Science 281:375-388(1998).  
CC -1- SIMILARITY: Belongs to the L29P family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001202; AAC65182.1; -.  
DR PIR; G71355; G71355.  
DR TIGR; TP0197; -.  
DR HAMAP; MF 00374; -; 1.  
DR InterPro; IPR001854; Ribosomal\_L29.  
DR TrEMBL; TIGR00012; L29; 1.  
DR PROSITE; PS00579; RIBOSOMAL\_L29; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 72 AA; 8654 MW; BF6163F9E244D63A CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 104 ILRRQI 109  
DB 45 ILRRQI 50  
  
RESULT 51  
Y056 NPVOP STANDARD; PRT; 82 AA.  
AC G0314;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 9.7 kDa protein (ORF60).  
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpNPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=164623;  
RN 1;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97271300; PubMed=9126251;  
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
RA Rohrmann G.F.;  
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear  
RT polyhedrosis virus genome.";  
RL Virology 229:381-399(1997).  
CC -1- SIMILARITY: TO CORRESPONDING ORF IN ACNPV.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U75930; AAC59059.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 82 AA; 9686 MW; 37AC49CD76E1673B CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 82;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 316 EQFADE 321  
DB 50 EQFADE 55  
  
RESULT 52  
Y46S SYN3 STANDARD; PRT; 83 AA.  
ID Y46S SYN3  
AC P73882;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein ssi0461.  
GN SSI0461.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN 1;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirose M., Sugita M., Saito M., Saito T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D90910; BAA17945.1; -.  
DR PIR; S75083; S75083.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 83 AA; 9471 MW; B8C3711A1F798D2C CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 66 VAQLAQ 71  
DB 32 VAQLAQ 37  
  
RESULT 53  
RL31 METJA STANDARD; PRT; 87 AA.  
ID RL31 METJA  
AC P34009;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L31e.  
GN RPL31E OR MJ0049.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX  MEDLINE=96337999; PubMed=8688087;
RA  Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA  Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA  Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA  Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA  Scott J.L., Geoghegan N.S.M., Weidman J.E., Fuhrmann J.L., Nguyen D.,
RA  Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA  Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA  Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT  "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT  jannaschii";
RL  Science 273:1058-1073(1996).
RC  -!- SIMILARITY: Belongs to the L31E family of ribosomal proteins.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U67463; AAB98030.1; -.
DR  FIRM; A64306; A64306.
DR  TIGR; M00049; -.
DR  HAMAP; MF_00410; -.
DR  InterPro; IPR000054; Ribosomal L31e.
DR  Pfam; PF01198; Ribosomal L31e; 1.
DR  PROSITE; PS01144; RIBOSOMAL_L31E; 1.
KW  Ribosomal protein; Complete proteome.
SQ  SEQUENCE 87 AA; 10205 MW; 74DEEF9900A1BF68 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 ARIVKI 20
DB 45 ARIVKI 50
|||||

RESULT 54
41_CHICK STANDARD; PRT; 90 AA.
AC P12264;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Erythroid protein 4.1 (Band 4.1) (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=87260822; PubMed=3474611;
RA Ngai J., Stack J.H., Moon R.T., Lazarides E.;
RT "Regulated expression of multiple chicken erythroid membrane skeletal
RT protein 4.1 variants is governed by differential RNA processing and
RT translational control.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4432-4436(1987).
CC -!- FUNCTION: Protein 4.1 is a major structural element of the
CC erythrocyte membrane skeleton. It plays a key role in regulating
CC membrane physical properties of mechanical stability and
CC deformability by stabilizing spectrin-actin interaction. Binds
CC with a high affinity to glycophorin and with lower affinity to
CC band III protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;

```

```

CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=P12264-1; Sequence=Displayed;
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M16962; AAA48762.1; -.
DR  PIR; A27056; A27056.
DR  InterPro; IPR000299; Band_4.1.
DR  Pfam; PF00373; Band_41; 1.
DR  PROSITE; PS00660; FERM_1; PARTIAL.
DR  PROSITE; PS00661; FERM_2; PARTIAL.
DR  PROSITE; PS00657; FERM_3; 1.
KW Structural protein; Alternative splicing; Cytoskeleton; Actin-binding;
KW Phosphorylation.
FT NON_TER 1 1 FERM.
FT DOMAIN <1 >90
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 10395 MW; 8938A0C88816604A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLV 80
DB 33 SSGLLV 38
|||||

RESULT 55
4BP3_MOUSE STANDARD; PRT; 101 AA.
ID 4BP3_MOUSE
AC Q80VV3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Eukaryotic translation initiation factor 4E-BP3.
DE (eIF4E-binding protein 3).
GN EIF4EBP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Poulin F., Sonenberg N.;
RT "Two overlapping reading frames in the second exon of the
RT translational inhibitor 4E-BP3.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., Mcowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale L.J., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Regulates eIF4E activity by preventing its assembly into  
CC the eIF4F complex (By similarity).  
CC -!- SUBUNIT: EIF4EBP3 interacts with EIF4E (By similarity).  
CC -!- PTM: Phosphorylated (By similarity).  
CC -!- SIMILARITY: Belongs to the eIF4E binding protein family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AY226182; AAC73448.1; -;  
DR EMBL; BC061242; AAH61242.1; -;  
KW Translation regulation; Protein synthesis inhibitor; Phosphorylation.  
SQ SEQUENCE 101 AA; 11018 MW; 78F92052696A9BD7 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 313 TDDEQF 318  
Db 92 TDDEQF 97  
|||||  
-----  
RESULT 56  
ID\_VNS7 CVCAE STANDARD; PRT; 101 AA.  
AC Q04703;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Nonstructural protein 7 (11 kDa protein) (X3 protein) (6A protein).  
GN NS7.  
OS Canine enteric coronavirus (strain X378) (CCoV) (CCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus.  
OX NCBI\_TaxID=33732;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93033103; PubMed=1329312;  
RA Vennema H., Rossen J.W.A., Wesseling J., Horzinek M.C.,  
RA Rottier P.J.M.;  
RT "Genomic organization and expression of the 3' end of the canine and  
RT feline enteric coronaviruses.";  
RL Virology 191:134-140(1992).  
CC -!- FUNCTION: MAY FUNCTION IN THE FORMATION OF MEMBRANE-BOUND  
CC REPLICATION COMPLEXES OR IN THE ASSEMBLY OF THE VIRUS.  
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).  
CC -!- SIMILARITY: Belongs to the coronavirus NS7 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X66717; CAA47247.1; -;  
DR F01; B44056; B44056.  
DR InterPro; IPR003449; Corona\_7.

DR Pfam; PF02398; Corona\_7; 1.  
DR ProDom; PD004753; Corona\_7; 1.  
KW Nonstructural protein; Membrane.  
SQ SEQUENCE 101 AA; 11491 MW; 1AB4641D94DB9BB1 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 78 LLVTLI 83  
Db 96 LLVTLI 101  
|||||  
-----  
RESULT 57  
ID\_Y18K MSVK STANDARD; PRT; 110 AA.  
AC P14988;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 12.7 kDa protein.  
OS Maize streak virus (Kenyan isolate) (MSV).  
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.  
OX NCBI\_TaxID=10822;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85037917; PubMed=6493977;  
RA Howell S.H.;  
RT "Physical structure and genetic organisation of the genome of maize  
RT streak virus (Kenyan isolate).";  
RL Nucleic Acids Res. 12:7359-7375(1984).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X01089; CAB37354.1; -;  
DR InterPro; IPR001191; Gemini\_ALI.  
DR ProDom; PD000736; Gemini\_ALI; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 110 AA; 12756 MW; FD028DB22508B037 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 110;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 61 PPEAV 66  
Db 100 PPEAV 105  
|||||  
-----  
RESULT 58  
ID\_YCX1 CHLPY STANDARD; PRT; 110 AA.  
AC P05720;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 12.7 kDa protein in 16S-23S DNA spacer.  
OS Chlorella pyrenoidosa.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
OC Chlorrellaceae; Chlorella.  
OX NCBI\_TaxID=3078;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86232622; PubMed=3714498;  
RA Yamada T., Shimaji M.;

RT "Peculiar feature of the organization of rRNA genes of the *Chlorella*  
 RL chloroplast DNA."; 14:3827-3939 (1986).  
 CC Nucleic Acids Res.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X03848; CAA27477.1; --  
 CC PIR; A24444; A24444.  
 KW Chloroplast; Hypothetical protein.  
 SQ SEQUENCE 110 AA; 12782 MW; 1F58A66055A1A377 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 300 KLIFFL 305  
 DB 56 KLIFFL 61  
 RESULT 59  
 SENA AFLCA  
 ID SENA AFLCA STANDARD; PRT; 113 AA.  
 AC P29233;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Sensorin A precursor.  
 GN PSC1.  
 OS Aplysia californica (California sea hare).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidea;  
 OC Aplysioidea; Aplysiidae; Aplysia.  
 OX NCBI\_TaxID=6500;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-54.  
 RC TISSUE=Pleural sensory cells;  
 RX MEDLINE=91227915; PubMed=1840700;  
 RA Brunet J.-F., Shapiro E., Foster S.A., Kandel E.R., Iino Y.;  
 RT "Identification of a peptide specific for Aplysia sensory neurons by  
 RL PCR-based differential screening."; Science 252:856-859(1991).  
 CC -!- FUNCTION: May function as an inhibitory cotransmitter acting in  
 CC conjunction with the fast excitatory transmitter released by  
 CC sensory neurons. The peptide selectively inhibits certain  
 CC postsynaptic cells probably by means of sensorin A release.  
 CC -!- SUBCELLULAR LOCATION: Throughout the neuronal cells (cell body,  
 CC axon and presynaptic terminals).  
 CC -!- TISSUE SPECIFICITY: Seems to be specific to the mechanosensory  
 CC neurons of the central nervous system.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X56770; CAA40089.1; --  
 CC PIR; S23653; S23653.  
 KW Signal; Neurone; Amidation; Cleavage on pair of basic residues.  
 FT SIGNAL 1 32 POTENTIAL.  
 FT PEPTIDE 33 54  
 FT PEPTIDE 46 54  
 FT PROPEP 58 113  
 FT MOD\_RES 54 54  
 FT AMIDATION (G-55 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 113 AA; 12711 MW; 74350F5154B49E1A CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 26 AILEKQ 31  
 DB 83 AILEKQ 88  
 RESULT 60  
 Y905 ARCFU  
 ID Y905 ARCFU STANDARD; PRT; 124 AA.  
 AC Q29357;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical UPF0292 protein AF0905.  
 GN AF0905.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.J., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyripides J.D.,  
 RA Reichmann D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
 RA Cotton P.W., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RL reducing archaeon *Archaeoglobus fulgidus*."; Nature 390:364-370(1997).  
 CC -!- SIMILARITY: Belongs to the UPF0292 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE001041; AAB90335.1; --  
 CC PIR; A69363; A69363.  
 CC TIGR; AF0905; --  
 CC HAMAP; MF 01095; --  
 CC InterPro; IPR006171; Toprim dom.  
 CC InterPro; IPR006154; Toprim sub.  
 CC Pfam; PF01751; Toprim; 1.  
 CC SMART; SM00493; TOPRIM; 1.  
 CC HYPOTHETICAL protein; Complete proteome.  
 SQ SEQUENCE 124 AA; 13948 MW; 55749C70EDCBC5B9 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 92 EGKKDV 97  
 DB 27 EGKKDV 32  
 RESULT 61

```

YC35 CYAPA
ID YC35_CYAPA STANDARD; PRT; 128 AA.
AC P48275;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 15.3 kDa protein ycf35.
GN YCF35.
OS Cyanophora paradoxa.
OG Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UTEX LB 555 / Pringsheim;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RT Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UTEX LB 555 / Pringsheim;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schlachter W.M., Chung S., Newmann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa: the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg (1997).
CC -1- SIMILARITY: Belongs to the ycf35 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30821; AAA81309.1; -
DR PIR; T06966; T06966.
DR CYanelle; Hypothetical protein.
SQ SEQUENCE 128 AA; 15317 MW; 0FEA01CDBC59BFA CRC64;

Query Match 1.8%; Score 6; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LIADLQ 87
DB 68 LIADLQ 73

RESULT 62
RS6E HALN1 STANDARD; PRT; 131 AA.
AC Q9HWJ5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S6e.
GN RPS6E OR VNG2514G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,

```

```

RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.B., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- SIMILARITY: Belongs to the S6E family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AR005129; AAG20576.1; -
DR PIR; D84401; D84401.
DR HAMAP; MF 00512; -; 1.
DR InterPro; IPR001377; Ribosomal_S6E.
DR Pfam; PF01092; Ribosomal_S6e; 1.
DR PROSITE; PS00578; RIBOSOMAL_S6E; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 131 AA; 13743 MW; 52FF77B556E2004B CRC64;

Query Match 1.8%; Score 6; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69
DB 103 EAVAQL 108

RESULT 63
RS8 CHLPN STANDARD; PRT; 133 AA.
AC Q3Z7S0; Q9JOG9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S8.
GN RPS8 OR RS8 OR CPN0634 OR CP0113 OR CPB0660.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

```

RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J133  
RL from Japan and CWL029 from USA.";  
RL Nucleic Acids Res. 28:2311-2314(2000).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=TW-183;  
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,  
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
RT other Chlamydia strains based on whole genome sequence analysis.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds  
CC directly to 16S rRNA central domain where it helps coordinate  
CC assembly of the platform of the 30S subunit (by similarity).  
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S5  
CC and S12 (by similarity).  
CC -!- SIMILARITY: Belongs to the 89P family of ribosomal proteins.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE001647; AAD18773.1; -.  
DR EMBL; AE002173; AAF37996.1; -.  
DR EMBL; AP002547; BAA98841.1; -.  
DR EMBL; AE017159; AAP98589.1; -.  
DR PIR; D72054; D72054.  
DR PIR; G86569; G86569.  
DR HSP; P56209; ISEI.  
DR TIGR; CP0113; -.  
DR HAMAP; MF 01302; -; 1.  
DR InterPro; IPR000630; Ribosomal\_S8.  
DR Pfam; PF00410; Ribosomal\_S8; 1.  
DR ProDom; PD01098; Ribosomal\_S8; 1.  
DR ProDom; PD00108; Ribosomal\_S8; 1.  
DR PROSITE; PS00053; RIBOSOMAL\_S8; 1.  
KW Ribosomal protein; rRNA-binding; rRNA-binding; Complete proteome.  
SQ SEQUENCE 133 AA; 15117 MW; 0EAE2623F8BE46A3 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 IVKILK 22  
DB 37 IVKILK 42  
|||||  
-----  
RESULT 64  
RL27 PYRST RL27 PYRST STANDARD; PRT; 134 AA.  
ID AC Q02984;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE 60S ribosomal protein L27.  
GN RPL27.  
OS Pyrobaculum stelleri (Chlamydomonads stellata).  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Spondylomoraceae; Pyrobaculum.  
OX NCBI\_TaxID=3064;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=94143474; PubMed=8310054;  
RA Wolf A.H., Kirsch M., Wiesner W.;  
RT "Nucleotide sequence of a cDNA encoding ribosomal protein L27 from  
RT Chlamydomonads stellata.";  
RL Plant Physiol. 101:1123-1123(1993).  
CC -!- SIMILARITY: Belongs to the L27E family of ribosomal proteins.

CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X68202; CAA48289.1; -.  
DR PIR; S26612; S26612.  
DR InterPro; IPR005824; KOW.  
DR InterPro; IPR006646; KOW sub.  
DR InterPro; IPR001141; Ribosomal\_L27e.  
DR Pfam; PF00467; KOW; 1.  
DR ProDom; PD01777; Ribosomal\_L27e; 1.  
DR ProDom; PD009396; Ribosomal\_L27e; 1.  
DR SMART; SM00739; KOW; 1.  
DR PROSITE; PS01107; RIBOSOMAL\_L27B; 1.  
KW Ribosomal protein.  
SQ SEQUENCE 134 AA; 15408 MW; 50BEC32EF9C259E CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 223 VTKRQS 228  
DB 53 VTKRQS 58  
|||||  
-----  
RESULT 65  
RR9 EUGR RR9 EUGR STANDARD; PRT; 134 AA.  
ID AC P32060;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chloroplast 30S ribosomal protein S9.  
GN RPS9.  
OS Euglena gracilis.  
OC Chloroplast.  
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.  
OX NCBI\_TaxID=3039;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=93347989; PubMed=8346031;  
RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,  
RA Orsat B., Spielmann A., Stutz E.;  
RT "Complete sequence of Euglena gracilis chloroplast DNA.";  
RL Nucleic Acids Res. 21:3537-3544(1993).  
CC -!- SIMILARITY: Belongs to the S9P family of ribosomal proteins.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z11874; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X70810; CAA50120.1; -.  
DR HAMAP; MF\_00532; -; 1.  
DR InterPro; IPR000754; Ribosomal\_S9.  
DR Pfam; PF00380; Ribosomal\_S9; 1.  
DR ProDom; PD001627; Ribosomal\_S9; 1.  
DR PROSITE; PS00360; RIBOSOMAL\_S9; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 134 AA; 15026 MW; 81C96A2E550E504F CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;



```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKIILF 161
Db 17 AKIILF 22

RESULT 66
IF2B_METH
ID IF2B_METH STANDARD; PRT; 135 AA.
AC 027797;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE Probable translation initiation factor 2 beta subunit (eIF-2-beta).
GN EIF2B OR WTH1769.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 178:7135-7155(1997).
CC -!- FUNCTION: eIF-2 functions in the early steps of protein synthesis
CC by forming a ternary complex with GTP and initiator tRNA (By
CC similarity).
CC -!- SUBUNIT: Heterotrimer composed of an alpha, a beta and a gamma
CC chain (By similarity).
CC -!- SIMILARITY: Belongs to the eIF-2-beta / eIF-5 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000932; AAB86235.1; -.
CC PIR; B69103; B69103.
CC HAVAP; MF_00232; -.
CC InterPro; IPR002735; eIF5_eIF2B.
CC InterPro; IPR004458; TIF_eIF2B.
CC Pfam; PF01873; eIF5_eIF2B; 1.
CC ProDom; PD004078; eIF5_eIF2B; 1.
CC SMART; SM00653; eIF2B_5; 1.
CC TIGRFAMs; TIGR00311; eIF-2beta; 1.
KW Initiation factor; Protein biosynthesis; Complete proteome.
SQ SEQUENCE 135 AA; 15567 MW; 98AF22F6966CA0AF CRC64;

Query Match 1.8%; Score 6; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
Db 3 DYKLL 8

RESULT 67
YJAV_BACSU
ID YJAV_BACSU STANDARD; PRT; 135 AA.

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
Db 3 DYKLL 8

RESULT 68
RR8_OENHO
ID -RR8_OENHO STANDARD; PRT; 138 AA.
AC Q9MTJ0;
DT 28-FEB-2003 (Rel. 41, Created)

```

```

AC 034959;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yJav.
GN YJAV OR BSU11290.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapich A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Roche E., Roche M., Rose M., Sadaie F.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokini A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z99109; CAB12970.1; -.
CC PIR; B69842; E69842.
CC Subtilist; BG13125; YJAV.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 135 AA; 16148 MW; 515D09CC417CFA7 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 QNYDTI 209
Db 66 QNYDTI 71

RESULT 68
RR8_OENHO
ID -RR8_OENHO STANDARD; PRT; 138 AA.
AC Q9MTJ0;
DT 28-FEB-2003 (Rel. 41, Created)

```

```
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DI Chloroplast 30S ribosomal protein S8.
GN RPS8.
OS Eucenothea hookeri (Hooker's evening primrose).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=85636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Johansen;
RX MEDLINE=20309318; PubMed=10852478;
RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
RA Chiu W.L., Sears B.;
RT "Complete nucleotide sequence of the Oenothera elata plastid
RT chromosome, representing plastome I of the five distinguishable
RT Eucenothea plastomes."
RL Mol. Genet. 263:581-585(2000).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA central domain where it helps coordinate
CC assembly of the platform of the 30S subunit (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the S8P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ274079; CAB67195.1; -.
CC HSP: P56209; ISEI.
CC HAMAP: MF_01302; -.
CC InterPro: IPR000630; Ribosomal_S8.
CC Pfam: PF00410; Ribosomal_S8_1.
CC ProDom: PD001098; Ribosomal_S8; 1.
CC PROSITE: PS00053; RIBOSOMAL_S8; 1.
CC Ribosomal protein; rRNA-binding; rRNA-binding; Chloroplast.
CC SEQUENCE 138 AA; 15843 MW; B047635ECE213F29 CRC64;
SQ
Query Match 1.8%; Score 6; DB 1; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.7e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 EIVKIL 21
DB 34 EIVKIL 39
RESULT 69
IF2B PYRAB STANDARD; PRT; 140 AA.
AC Q9UVF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable translation initiation factor 2 beta subunit (eIF-2-beta).
GN EIF2B OR PYRAB14410 OR PAB0959.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GE5 / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Pech O., Prieur D., Querellou J., Ripp R., Thiery J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RA "An integrated analysis of the genome of the hyperthermophilic
RA archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
CC -!- FUNCTION: eIF-2 functions in the early steps of protein synthesis
CC by forming a ternary complex with GTP and initiator tRNA (By
CC similarity).
CC -!- SUBUNIT: Heterotrimer composed of an alpha, a beta and a gamma
CC chain (By similarity).
CC -!- SIMILARITY: Belongs to the eIF-2-beta / eIF-5 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ248287; CAB50346.1; -.
CC FIR: EY5056; EY5056.
CC HAMAP: MF_00232; -.
CC InterPro: IPR002735; eIF5_eIF2B.
CC InterPro: IPR004458; TIF_aIF-2beta.
CC Pfam: PF01873; eIF5_eIF2B; 1.
CC ProDom: PD004078; eIF5_eIF2B; 1.
CC SMART: SM00653; eIF2B_5; 1.
CC TIGRFAMs: TIGR00311; aIF-2beta; 1.
CC Initiation factor; Protein biosynthesis; Complete proteome.
CC SEQUENCE 140 AA; 16247 MW; 9D40F2C556DB539A CRC64;
SQ
Query Match 1.8%; Score 6; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 212 DYKLL 217
DB 7 DYKLL 12
RESULT 70
IF2B PYRFU STANDARD; PRT; 140 AA.
AC Q8UJ15;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable translation initiation factor 2 beta subunit (eIF-2-beta).
GN EIF2B OR PF0481.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: eIF-2 functions in the early steps of protein synthesis
CC by forming a ternary complex with GTP and initiator tRNA (By
CC similarity).
CC -!- SUBUNIT: Heterotrimer composed of an alpha, a beta and a gamma
CC chain (By similarity).
CC -!- SIMILARITY: Belongs to the eIF-2-beta / eIF-5 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
-----
DR EMBL; AE010171; AAL80605.1; -.
DR HAMAP; MF_00232; -.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR004458; TIF_aIF-2beta.
DR Pfam; PF01873; eIF5_eIF2B; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR TIGRFAMs; TIGR00311; aIF-2beta; 1.
KW Initiation factor; Protein biosynthesis; Complet  proteome.
SQ SEQUENCE 140 AA; 16234 MW; 771153C98A871359 CRC64;

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 140;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
DB 7 DYKLL 12

RESULT 71
IF2B_PYRHO STANDARD; PRT; 140 AA.
AC O58312;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable translation initiation factor 2 beta subunit (eIF-2-beta).
GN EIF2B OR PH0605.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Hirakawa H., Haikawa Y., Hino Y.,
Yamamoto S., Sekine M., Baba S.-I., Kobugi H., Hosoyama A., Nagai Y.,
Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
Masuchi Y., Shizuwa H., Kikuchi H.
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -!- FUNCTION: eIF-2 functions in the early steps of protein synthesis
by forming a ternary complex with GTP and initiator tRNA (By
similarity).
CC -!- SUBUNIT: Heterotrimer composed of an alpha, a beta and a gamma
chain (By similarity).
CC -!- SIMILARITY: Belongs to the eIF-2-beta / eIF-5 family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; APO00003; BAA29694.1; -.
PIR; D71104; D71104.
DR HAMAP; MF_00232; -.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR004458; TIF_aIF-2beta.
DR Pfam; PF01873; eIF5_eIF2B; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR TIGRFAMs; TIGR00311; aIF-2beta; 1.
KW Initiation factor; Protein biosynthesis; Complete proteome.
SQ SEQUENCE 140 AA; 16247 MW; 9D40F2C1428A129A CRC64;
```

```
-----
Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 140;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
DB 7 DYKLL 12

RESULT 72
U426_HSVMG STANDARD; PRT; 142 AA.
AC Q05104;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 15.7 kDa protein.
GN U426.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118245; PubMed=1282282;
RA Sakaguchi M., Urakawa T., Hirayama Y., Miki N., Yamamoto M.,
Hirai K.;
RT "Sequence determination and genetic content of an 8.9-kb restriction
fragment in the short unique region and the internal inverted repeat
of Marek's disease virus type 1 DNA.";
RL Virus Genes 6:365-378(1992).
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M80595; AAB59898.1; -.
KW Hypothetical protein.
SQ SEQUENCE 142 AA; 15662 MW; 76D137DF02735E26 CRC64;

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 142;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLG 233
DB 25 SLKLG 30

RESULT 73
HS16_CAEEL STANDARD; PRT; 143 AA.
AC P06581; P02514;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock protein HSP16-41.
GN HSP16-41 OR Y46H3A.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86304344; PubMed=3017958;
RA Jones D., Russnak R.H., Kay R.J., Candido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
Caenorhabditis elegans that is flanked by repetitive elements.";
```

```
RL J. Biol. Chem. 261:12006-12015(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hristol N2;
RA Stoneking T., Wohlmann P., Lennox S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterson R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 47-143 FROM N.A.
RX MEDLINE=83220736; PubMed=6190129;
RA Rusanak R.H., Jones D., Candido E.P.M.;
RT "Cloning and analysis of cDNA sequences coding for two 16 kilodalton
RT heat shock proteins (hsp) in Caenorhabditis elegans: homology with
RT the small hsp of Drosophila.";
RL Nucleic Acids Res. 11:3187-3205(1983).
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14334; AAA28070.1; ALT_SEQ.
DR EMBL; AC006774; AAF60616.1; -.
DR EMBL; X01577; CAA25732.1; -.
DR PIR; A25199; HHKW41
DR WormPep; Y46H3A.2; CE22003.
DR InterPro; IPR001436; Crystallin_alpha.
DR InterPro; IPR002068; Hsp20.
DR InterPro; IPR008978; HSP20_chap.
DR Pfam; PF00011; HSP20; 1.
DR PRINTS; PR00299; ACRYSTALLIN.
DR PROSITE; PS01031; HSP20; 1.
DR Heat shock; Multigene family.
KW Heat shock; 143 AA; 16252 MW; C1D0F59D26E36C24 CRC64;
SQ SEQUENCE 143 AA; 16252 MW;

Query Match 1.8%; Score 6; DB 1; Length 143;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPENLK 256
DB 64 KPENLK 69

RESULT 74
RL9_UREPA
ID RL9_UREPA STANDARD; PRT; 145 AA.
AC Q9PPT9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L9.
GN RPLI OR RPL9 OR U551.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Iefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum.";
RL Nature 407:757-762(2000).

J. Biol. Chem. 261:12006-12015(1986).
-!- FUNCTION: Binds to the 23S rRNA (By similarity).
-!- SIMILARITY: Belongs to the L9p family of ribosomal proteins.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE002153; AAF30964.1; -.
DR HSP; P02417; 1DIV.
DR HAMAP; MF_00503; 1.
DR InterPro; IPR009027; L9_N_like.
DR InterPro; IPR000244; Ribosomal_L9.
DR Pfam; PF03948; Ribosomal_L9_C; 1.
DR Pfam; PF01281; Ribosomal_L9_N; 1.
DR TIGRFAMs; TIGR00158; L9_1.
DR PROSITE; PS00653; RIBOSOMAL_L9; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 145 AA; 16420 MW; 3B5E203806EDF15F CRC64;

Query Match 1.8%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295
DB 125 IVEILL 130

RESULT 75
TAAL_VACCC
ID TAAL_VACCC STANDARD; PRT; 150 AA.
AC P20982;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-activator protein A1.
GN All.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -!- FUNCTION: General late promoter trans-activator.
CC -!- DEVELOPMENTAL STAGE: Intermediate stages of infection.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35027; AAA48116.1; -.
DR PIR; C42517; C42517.
DR InterPro; IPR004975; Pox_TAAL1.
DR Pfam; PF03295; Pox_TAAL1; 1.
```



```

Query Match      1.8%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLRDK 265
DB 121 NLRDK 126

RESULT 78
RUVX ANASP
ID RUVX ANASP STANDARD; PRT; 151 AA.
AC Q8YS15;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative Holliday junction resolvase (EC 3.1.-.-).
GN ALK3100.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kameko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: Could be a nuclease that resolves Holliday junction
CC intermediates in genetic recombination.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the yggF Hcr family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AP003591; BAB74799.1; -.
DR PIR; AE2193; AE2193.
DR HAMAP; MF_00651; -.
DR InterPro; IPR005227; Cons_hypoth250.
DR Pfam; PF03652; UPF0081; 1.
DR SMART; SM00732; YggF; 1.
DR TIGRFAMs; TIGR00250; TIGR00250; 1.
KW Hydrolase; Nuclease; DNA repair; DNA recombination; Complete proteome.
SQ SEQUENCE 151 AA; 16799 MW; 0F6618C1E0381E68 CRC64;

Query Match      1.8%; Score 6; DB 1; Length 151;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 MTKYIS 250
DB 1 MTKYIS 6

RESULT 79
SSRP HELPJ
ID SSRP HELPJ STANDARD; PRT; 152 AA.
AC Q9ZJH2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SsrA-binding protein.

```

```

GN SMPB OR JHP1337.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown B.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- FUNCTION: Binds specifically to the ssrA RNA (tmRNA) and is
CC required for stable association of ssrA with ribosomes (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the smpB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001556; AAD06913.1; -.
DR PIR; G71820; G71820.
DR HAMAP; MF_00023; -.
DR InterPro; IPR000037; SmpB.
DR Pfam; PF01669; SmpB; 1.
DR ProDom; PD004488; SmpB; 1.
DR TIGRFAMs; TIGR00086; smpB; 1.
DR PROSITE; PS01317; SSRP; 1.
KW RNA-binding; Complete proteome.
SQ SEQUENCE 152 AA; 17814 MW; 2A07C14BB5E2C364 CRC64;

Query Match      1.8%; Score 6; DB 1; Length 152;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KRQSLK 230
DB 129 KRQSLK 134

RESULT 80
SSRP HELPJ
ID SSRP HELPJ STANDARD; PRT; 152 AA.
AC O25985;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SsrA-binding protein.
GN SMPB OR HP1444.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Klenk H.-P., Gill S., Dougherty B.A.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Kirkness E.F., Peterson S.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Glodek A.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

```





```

KW T-cell.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 155 INTERLEUKIN-2.
FT DISULFID 79 127 BY SIMILARITY.
FT CARBOHYD 23 23 O-LINKED (GALNAC... ) (BY SIMILARITY).
FT CONFLICT 66 66 V -> A (IN REF. 2).
SQ SEQUENCE 155 AA; 17627 MW; 816667DFA052EDF CRC64;

Query Match 1.8%; Score 6; DB 1; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLXL 257
DB 51 PENLXL 56
|||||

RESULT 83
IL2_BUBBU
ID IL2_BUBBU STANDARD; PRT; 155 AA.
AC Q95KP3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Interleukin-2 precursor (IL-2) (T-cell growth factor) (TCGF).
GN IL2.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22117321; PubMed=12121282;
RA Sreekumar E., Premraj A., Saravanakumar M., Rasool T.J.;
RT "Buffalo (Bubalus bubalis) interleukin-2: sequence analysis reveals
RT high nucleotide and amino acid identity with interleukin-2 of cattle
RT and other ruminants."
RL Eur. J. Immunogenet. 29:341-345(2002).
CC -1- FUNCTION: Produced by T-cells in response to antigenic or
CC mitogenic stimulation, this protein is required for T-cell
CC proliferation and other activities crucial to regulation of the
CC immune response. Can stimulate B cells, monocytes, lymphokine-
CC activated killer cells, natural killer cells, and glioma cells (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the IL-2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF363786; AAK50039.1; -.
CC DR InterPro; IPR000779; Interleukin-2.
CC DR Pfam; PF00715; IL2; 1.
CC DR PRINTS; PR00265; INTERLEUKIN2.
CC DR ProDom; PD003649; Interleukin-2; 1.
CC DR SMART; SM00189; IL2; 1.
CC DR PROSITE; PS00424; INTERLEUKIN 2; 1.
KW Cytokine; Glycoprotein; Immune response; Signal; Growth factor;
KW T-cell.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 155 INTERLEUKIN-2.
FT DISULFID 79 127 BY SIMILARITY.
FT CARBOHYD 23 23 O-LINKED (GALNAC... ) (BY SIMILARITY).
SQ SEQUENCE 155 AA; 17541 MW; 1753173FE619946B CRC64;

Query Match 1.8%; Score 6; DB 1; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLXL 257
DB 51 PENLXL 56
|||||

RESULT 84
IL2_SHEEP
ID IL2_SHEEP STANDARD; PRT; 155 AA.
AC P19I14; Q95MP4;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Interleukin-2 precursor (IL-2) (T-cell growth factor) (TCGF).
GN IL2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91016933; PubMed=2216781;
RA Goodall J.C., Emery D.C., Perry A.C.F., English L.S., Hall L.;
RT "cDNA cloning of ovine interleukin 2 by PCR."
RL Nucleic Acids Res. 18:5883-5883(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088336; PubMed=2263496;
RA Seow H.F., Rothel J.S., Radford A.J., Wood P.R.;
RT "The molecular cloning of ovine interleukin 2 gene by the polymerase
RT chain reaction."
RL Nucleic Acids Res. 18:7175-7175(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20500438; PubMed=11048943;
RA Luehken G., Hiendler S., Prinzenberg E.M., Erhardt G.;
RT "A single-strand conformation polymorphism in the ovine interleukin-2
RT (IL-2) gene."
RL J. Anim. Sci. 78:2754-2755(2000).
RN [4]
RP SEQUENCE OF 21-153 FROM N.A.
RX MEDLINE=9537627; PubMed=7543777;
RA Bujdosó R., Williamson M.L., Roy D., Hunt P., Blacklaws B., Sargan D.,
RA McConnell I.;
RT "Molecular cloning and expression of DNA encoding ovine interleukin
RT 2."
CY Cytokine 7:223-231(1995).
CC -1- FUNCTION: Produced by T-cells in response to antigenic or
CC mitogenic stimulation, this protein is required for T-cell
CC proliferation and other activities crucial to regulation of the
CC immune response. Can stimulate B cells, monocytes, lymphokine-
CC activated killer cells, natural killer cells, and glioma cells.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the IL-2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X53934; CAA37881.1; -.
CC DR EMBL; X55641; CAA39165.1; -.
CC DR EMBL; AF287479; AAK69554.1; -.
CC DR EMBL; X60148; CAA42722.1; -.
CC DR EMBL; A19169; CAA01448.1; -.
CC DR PIR; S11488; S11488.
CC DR HSSP; P01585; 3INK.
CC DR InterPro; IPR000779; Interleukin-2.

```

```

DR Pfam: PF00715; IL2; 1.
DR PRINTS; PRO0265; INTERLEUKIN2.
DR ProDom; PD003649; Interleukin-2; 1.
DR SMART; SM00189; IL2; 1.
DR PROSITE; PS00424; INTERLEUKIN 2; 1.
KW Cytokine; Glycoprotein; Immune response; Signal; Growth factor;
KW T-cell.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 155 INTERLEUKIN-2.
FT CARBOHYD 23 23 O-LINKED (GALNAC. .) (BY SIMILARITY).
FT DISULFID 79 127 BY SIMILARITY.
FT CONFLICT 6 L -> P (IN REF. 1).
SQ SEQUENCE 155 AA; 17678 MW; 3F76C6CBB7B289C5 CRC64;

Query Match
Best Local Similarity 100.0%; DB 1; Length 155;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLKL 257
DB 51 PENLKL 56

RESULT 85
RL10_MYCPE
AC O8EV70; STANDARD; PRT; 161 AA.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR MYPE5740.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.,
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans."
RL Nucleic Acids Res. 30:5293-5300(2002).
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004172; BAC4364.1; ALT_INIT.
CC HAMAP; MF_00362; 1.
CC InterPro; IPR001790; Ribosomal_L10.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC Pfam; PF00466; Ribosomal_L10; 1.
CC PROSITE; PS01109; RIBOSOMAL_L10; FALSE NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 161 AA; 17949 MW; 1D95D1C29B5E87D3 CRC64;

Query Match
Best Local Similarity 100.0%; DB 1; Length 161;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107
DB 57 NNILRR 62

RESULT 86
RL10_MYCPE
AC P75240; STANDARD; PRT; 161 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10.
GN RPLJ OR MPNS38 OR MP304.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himelreich R., Hilbert H., Piagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000028; AAB95952.1; -.
CC PIR; S73630; S73630.
CC HAMAP; MF_00362; 1.
CC InterPro; IPR001790; Ribosomal_L10.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC Pfam; PF00466; Ribosomal_L10; 1.
CC PROSITE; PS01109; RIBOSOMAL_L10; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 161 AA; 17618 MW; 4A7C387BBA49D446 CRC64;

Query Match
Best Local Similarity 100.0%; DB 1; Length 161;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107
DB 56 NNILRR 61

RESULT 87
IL2_CEREL
ID IL2_CEREL STANDARD; PRT; 162 AA.
AC PS1747;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interleukin-2 precursor (IL-2) (T-cell growth factor) (TCGF).
GN IL2.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860;
RN [1]
RP SEQUENCE FROM N.A.
RA Lockhart E.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Produced by T-cells in response to antigenic or
CC mitogenic stimulation, this protein is required for T-cell
CC proliferation and other activities crucial to regulation of the
CC immune response. Can stimulate B cells, monocytes, lymphokine-
CC activated killer cells, natural killer cells, and glioma cells (By
CC similarity).

```



```
RESULT 90
RS16_CORGL STANDARD; PRT; 165 AA.
ID RS16_CORGL
AC QNNX3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S16.
GN RSP OR CGL2054.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the S16P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005280; BAB99447.1; -.
DR HAMAP; MF 00385; -.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.
DR ProDom; PD003791; Ribosomal_S16; 1.
DR TIGRFAMs; TIGR00002; S16; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 165 AA; 17837 MW; 61DD81961BC30846 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 PTEAVA 67
DB 66 PTEAVA 71

RESULT 91
GVA2_STRCO STANDARD; PRT; 170 AA.
ID GVA2_STRCO
AC Q9RJB4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable gas vesicle structural protein 2 (GVP).
GN GVA2 OR SCO0650 OR SCP91.10.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
```

```
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- FUNCTION: Gas vesicles are small, hollow, gas filled protein
CC structures that are found in several microbial planktonic
CC microorganisms. They allow the positioning of the organism at
CC the favorable depth for growth. GvpA type proteins form the
CC essential core of the structure.
CC -1- SUBCELLULAR LOCATION: Gas vesicle membrane.
CC -1- SIMILARITY: Belongs to the gas vesicle protein type A family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL939106; CAB61167.1; -.
DR HAMAP; MF 00576; -.
DR InterPro; IPR000638; Gas_vesicle.
DR Pfam; PF00741; Gas_vesicle; 1.
DR ProDom; PD003598; Gas_vesicle; 1.
DR PROSITE; PS00234; GAS_VESICLE_A_1; 1.
DR PROSITE; PS00669; GAS_VESICLE_A_2; FALSE_NEG.
KW Gas vesicle; Complete proteome.
SQ SEQUENCE 170 AA; 19422 MW; 57C97217BE59ACDC CRC64;

Query Match 1.8%; Score 6; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 ELILDR 239
DB 22 ELILDR 27

RESULT 92
SODC_CAUCR STANDARD; PRT; 174 AA.
ID SODC_CAUCR
AC P20379;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1).
GN SODC OR CC1579.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=90264275; PubMed=2345128;
RA Steinman H.M., Ely B.;
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

"Cu-Zn superoxide dismutase of Caulobacter crescentus: cloning,
sequencing, and mapping of the gene and periplasmic location of the
enzyme.";
J. Bacteriol. 172:2901-2910(1990).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
```

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=82265686; PubMed=7050107;  
 RA Steinman H.M.;  
 RT "Copper-zinc superoxide dismutase from *Caulobacter crescentus* CB15. A  
 novel bacteriophage protein form of the enzyme.";  
 RL J. Biol. Chem. 257:10283-10293(1982).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems (By similarity).  
 CC May function against extracytoplasmic toxic oxygen species.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COPACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
 similarity).  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic.  
 CC -1- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M55259; AAA23054.1; -;  
 DR EMBL; AE005832; AAK23558.1; -;  
 DR PIR; A35383; A35383.  
 DR HSP; P00446; IYAI.  
 DR TIGR; CC1579; -;  
 DR InterPro; IPR001424; SOD\_CU\_ZN.  
 DR Pfam; PF00080; sodcu; 1.  
 DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
 DR PROSITE; PS00087; SOD\_CU\_ZN 1; 1.  
 DR PROSITE; PS00332; SOD\_CU\_ZN 2; 1.  
 KW Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc; Periplasmic;  
 KW Signal; Complete proteome.  
 FT SIGNAL 1 23  
 FT CHAIN 24 174 SUPEROXIDE DISMUTASE [CU-ZN].  
 FT METAL 68 68 COPPER (BY SIMILARITY).  
 FT METAL 70 70 COPPER (BY SIMILARITY).  
 FT METAL 86 86 COPPER AND ZINC (BY SIMILARITY).  
 FT METAL 95 95 ZINC (BY SIMILARITY).  
 FT METAL 104 104 ZINC (BY SIMILARITY).  
 FT METAL 107 107 ZINC (BY SIMILARITY).  
 FT METAL 150 150 COPPER (BY SIMILARITY).  
 FT DISULFID 75 170 BY SIMILARITY.  
 SQ SEQUENCE 174 AA; 17100 MW; F3B3C79EF3E3642C CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 285 HKTOPI 290  
 Db 156 HKTOPI 161  
 RESULT 93  
 CYB\_NYCHU STANDARD; PRT; 176 AA.  
 AC Q36572;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cytochrome b (Fragment).  
 GN MTYB OR COB OR CYTB.  
 OS *Nycticeius humeralis* (Evening bat).  
 OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Vespertilionidae;  
 OC *Nycticeius*.  
 OX NCBI\_TaxID=27670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Isolate LSUMZ 23913; TISSUE=Kidney, and Liver;  
 RA Sudman P.D., Barkley L.J., Hafner M.S.;  
 RT "Familial affinity of *Tomopeas ravus* (Chiroptera) based on protein  
 electrophoretic and cytochrome b sequence data.";  
 RL J. Mammal. 75:365-377(1994).  
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
 complex (complex III or cytochrome b-c1 complex), which is a  
 respiratory chain that generates an electrochemical potential  
 coupled to ATP synthesis (By similarity).  
 CC -1- COPACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or  
 B562) is low-potential and absorbs at about 562, and heme 2 (or BH  
 or B566) is high-potential and absorbs at about 566 (By  
 similarity).  
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,  
 cytochrome c1 and the Rieske protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome b family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L19727; AAA17773.1; -;  
 DR InterPro; IPR005798; Cytb\_b6\_C.  
 DR InterPro; IPR005797; Cytb\_b6\_N.  
 DR Pfam; PF00033; cytochrome\_b\_N; 1.  
 DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
 DR PROSITE; PS00193; CYTOCHROME\_B\_QO; PARTIAL.  
 KW Mitochondrion; Electron transport; Respiratory chain; Transmembrane;  
 KW Heme.  
 FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).  
 FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).  
 FT NON TER 176 176  
 SQ SEQUENCE 176 AA; 19690 MW; 0D271F45A388DF5F CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 154 PLAKII 159  
 Db 9 PLAKII 14  
 RESULT 94  
 RPOE\_STAAM STANDARD; PRT; 176 AA.  
 AC Q99SD0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable DNA-directed RNA polymerase delta subunit (RNAP delta  
 factor).  
 GN RPOE OR SAV2128 OR SA1930 OR MW2052.  
 OS *Staphylococcus aureus* (strain Mu50 / ATCC 700699),  
 OS *Staphylococcus aureus* (strain N315), and  
 OS *Staphylococcus aureus* (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.  
 OX NCBI\_TaxID=158878, 158879, 196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mu50 / ATCC 700699, and N315;  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Kazutani-Di Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaico C.,  
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
RT "Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*."  
RL Lancet 357:1225-1240 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiramatsu K.,  
RT "Genome and virulence determinants of high virulence community-acquired MRSA."  
RL Lancet 359:1819-1827 (2002).  
CC -!- FUNCTION: Participates in both the initiation and recycling phases of transcription. In the presence of the delta subunit, RNAP displays an increased specificity of transcription, a decreased affinity for nucleic acids, and an increased efficiency of RNA synthesis because of enhanced recycling (By similarity).  
CC -!- SUBUNIT: RNAP is composed of a core of 2 alpha, a beta and a beta' subunits. The core is associated with a delta subunit and one of several sigma factors (By similarity).  
CC -!- SIMILARITY: Belongs to the rpoE family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AP003364; BAB58290.1; -;  
CC EMBL; AP003136; BAB43214.1; -;  
CC EMBL; AP004829; BAB95917.1; -;  
CC PIR; E90006; E90006.  
CC HAMAP; MF 00357; -; 1.  
CC InterPro; IPR007759; RNA\_pol\_delta.  
CC Pfam; PF05066; RNA\_pol\_delta\_1.  
CC KW Domain-directed RNA polymerase; Transcription; Complete proteome.  
FT DOMAIN 99 176 ASP/GLU-RICH (ACIDIC).  
SQ SEQUENCE 176 AA; 20881 MW; D7A72313DA512D89 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 176;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 229 LKLLGE 234  
Db 111 LKLLGE 116  
|||||  
RESULT 95  
HPRT\_ECOLI STANDARD; PRT; 178 AA.  
AC P36766;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) (HPRT).  
GN HPRT OR B0125 OR C0154 OR SF0122 OR S0124.  
OS Escherichia coli,  
OS Escherichia coli O6, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217932, 623;  
RN [1]

RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / W3110;  
RX MEDLINE=94261430; PubMed=8202364;  
RA Fujita N., Mori H., Yura T., Ishihama A.;  
RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the 2.4-4.1 min (110,917-193,643 bp) region."  
RL Nucleic Acids Res. 22:1637-1639 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
RT "The complete genome sequence of *Escherichia coli* K-12."  
RL Science 277:1453-1474 (1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.-A., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=22272406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J., Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S., Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y., Yu J.;  
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157."  
RL Nucleic Acids Res. 30:4432-4441 (2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W., Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A., Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S., Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T."  
RL Infect. Immun. 71:2775-2786 (2003).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=22065829; PubMed=12070315;  
RA Guddat L.W., Vos S., Martin J.L., Keough D.T., de Jersey J.;  
RT "Crystal structures of free, IMP-, and GMP-bound *Escherichia coli* hypoxanthine phosphoribosyltransferase."  
RL Protein Sci. 11:1626-1638 (2002).  
CC -!- FUNCTION: THIS ENZYME ACTS EXCLUSIVELY HYPOXANTHINE; IT DOES NOT ACT ON GUANINE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-alpha-D-ribose 1-diphosphate.  
CC -!- COFACTOR: Binds 2 magnesium ions per subunit. One of the ions does not make direct protein contacts (By similarity).  
CC -!- PATHWAY: Purine salvage.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the purine/pyrimidine phosphoribosyltransferase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
LT2";  
RL Nature 413:852-856 (2001).  
CC -1- FUNCTION: THIS ENZYME ACTS EXCLUSIVELY HYPOXANTHINE; IT DOES NOT  
CC ACT ON GUANINE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-  
CC alpha-D-ribose 1-diphosphate.  
CC -1- COFACTOR: Binds 2 magnesium ions per subunit. One of the ions does



RA Cooke R., Laudie M., Berger-Lilauro C., Purnelle B., Masuy D.,  
RA de Haan M., Waars A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.B., Tallon L.J., Jenkins J.,  
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
RA Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,  
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,  
RA Kiyokawa S., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakayama C., Nakazaki N., Shino S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,  
RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
RA thaliana.",  
RL Nature 408:820-822(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.A.,  
PT "Full-length cDNA from Arabidopsis thaliana";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=22954850; PubMed=14593172;  
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
RA Chao Q., Choy N., Ejima A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,  
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
RA Satou M., Tamez R., Vayaberg M., Wallender E.K., Wong C., Yamamura Y.,  
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,  
RT "Empirical analysis of transcriptional activity in the Arabidopsis  
RT genome";  
RL Science 302:842-846(2003).  
CC -!- FUNCTION: Participates in various redox reactions through the  
CC reversible oxidation of the active center dithiol to a disulfide.  
CC The F form is known to activate a number of enzymes of the  
CC photosynthetic carbon cycle (By similarity).  
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT P-TYPE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF144385; AAD35003.1; -  
CC EMBL: AC018363; AAF26987.1; -  
CC EMBL: AY084778; AAM61345.1; -  
CC EMBL: AY065391; AAL38832.1; -  
CC EMBL: AY096721; AAM20355.1; -  
CC HSP: P10599; 1A1U.  
CC InterPro: IPR006662; Thiorred.  
CC InterPro: IPR006663; Thiorredox\_dom2.  
CC Pfam: PF00085; thiorred; 1.  
CC PRINTS: PR00421; THIOREDOXIN.  
CC PROSITE: PS00194; THIOREDOXIN; 1.  
CC Redox-active center; Electron transport; Chloroplast; Transit peptide;  
CC Multigene family.  
KW TRANSIT 1 57 CHLOROPLAST (POTENTIAL).  
FT CHAIN 58 178 THIOREDOXIN F-TYPE 1.  
FT DISULFID 99 102 REDOX-ACTIVE (BY SIMILARITY).

FT CONFLICT 134 134 A -> P (IN REF. 1).  
SQ SEQUENCE 178 AA; 19325 MW; 7B4E6CF8F2714ED CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 178;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 KILKDN 24  
|||||  
DB 146 KILKDN 151  
RESULT 98  
Y969 VIBPA STANDARD; PRT; 179 AA.  
ID Y969 VIBPA  
AC Q87R29;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein VP0969.  
GN VP0969.  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
CC NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshina K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yaenaga T., Honda T., Shinagawa H., Hattori M., Iida T.,  
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
RT distinct from that of V. cholerae";  
RL Lancet 361:743-749(2003).  
CC -!- SIMILARITY: Belongs to the UPF0227 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AP005076; BACS59232.1; -  
CC HAMAP: MF\_01047; -; 1.  
CC InterPro: IPR008886; UPF0227.  
CC Pfam: PF05728; UPF0227; 1.  
CC Hypothetical protein; Complete proteome.  
SQ SEQUENCE 179 AA; 20742 MW; F016FB0E5F63AAD7 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 179;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 46 LQAMKE 51  
|||||  
DB 170 LQAMKE 175  
RESULT 99  
SYDP SALT1  
ID SYDP SALT1 STANDARD; PRT; 181 AA.  
AC P60085; Q8XFC9;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Syd protein.  
GN SYD OR STV3106 OR T2875.  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.

```

OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.B., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RN Nature 413:848-852(2001).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Interacts with the secY protein in vivo. May bind
CC preferentially to an uncomplexed state of secY, thus functioning
CC either as a chelating agent for excess secY in the cell or as a
CC regulatory factor that negatively controls the translocase
CC function (By similarity).
CC -!- SUBCELLULAR LOCATION: Loosely associated with the cytoplasmic side
CC of the inner membrane, probably via secY (By similarity).
CC -!- SIMILARITY: Belongs to the Syd family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE016843; AA070431.1; -.
CC EMBL; AL627276; CAD06080.1; -.
CC HAMAP; MF_01104; -; 1.
CC Complete proteome.
CC SEQUENCE 181 AA; 20518 MW; B5A36E754F43C8D2 CRC64;
CC
CC Query Match 1.8%; Score 6; DB 1; Length 181;
CC Best Local Similarity 100.0%; Pred. No. 2.1e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 317 QFADEK 322
CC Db 92 QFADEK 97
CC
CC RESULT 100
CC SYDP SALTY STANDARD; PRT; 181 AA.
CC AC P60084; Q8XFC9;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Syd protein.
CC GN SYD OR STM2967.
CC OS Salmonella typhimurium.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Salmonella.
CC OX NCBI_TaxID=602;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

```

```

RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Interacts with the secY protein in vivo. May bind
CC preferentially to an uncomplexed state of secY, thus functioning
CC either as a chelating agent for excess secY in the cell or as a
CC regulatory factor that negatively controls the translocase
CC function (By similarity).
CC -!- SUBCELLULAR LOCATION: Loosely associated with the cytoplasmic side
CC of the inner membrane, probably via secY (By similarity).
CC -!- SIMILARITY: Belongs to the Syd family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE008836; AAL21846.1; -.
CC STyGene; SG??777; Syd.
CC HAMAP; MF_01104; -; 1.
CC Complete proteome.
CC SEQUENCE 181 AA; 20518 MW; B5A36E754F43C8D2 CRC64;
CC
CC Query Match 1.8%; Score 6; DB 1; Length 181;
CC Best Local Similarity 100.0%; Pred. No. 2.1e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 317 QFADEK 322
CC Db 92 QFADEK 97
CC
CC RESULT 101
CC YHCE ECOLI
CC ID YHCE ECOLI STANDARD; PRT; 181 AA.
CC AC P45421; P76674;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein Yhce.
CC GN YHCE OR B3217.
CC OS Escherichia coli.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=562;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=K12 / MG1655;
CC RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
CC [2]
CC CONCEPTUAL TRANSLATION.
CC RU Rudd K.E.;
CC UNpublished observations (APR-1995).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THIS PROTEIN IS
CC TRUNCATED BY AN ISS ELEMENT WHICH IS INSERTED BETWEEN POSITION 123
CC AND 124.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U18997; AAA58019.1; ALT\_SEQ.  
DR EMBL; AE000401; AAC76249.1; ALT\_SEQ.  
DR EcoGene; EG12811; yhcE.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 181 AA; 19152 MW; 6C38BD142BD90B5A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 181;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 IASDAF 184  
Db 92 IASDAF 97

## RESULT 102

ID THIF BRANA STANDARD; PRT; 182 AA.  
AC O48897;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Thioresoxin F-type, chloroplast precursor (TRX-F).  
GN TRXF.

OS Brassica napus (rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3708;  
[1]

RP SEQUENCE FROM N.A.  
RA Mora-Garcia S.E.F., Rodriguez-Suarez R.J., Wolosinski R.A.;  
RA Submitted (AUG-1997) to the EMBL/GenBank/DBS databases.  
CC -!- FUNCTION: Participates in various redox reactions through the  
CC reversible oxidation of the active center dithiol to a disulfide.  
CC The F form is known to activate a number of enzymes of the  
CC photosynthetic carbon cycle (By similarity).  
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT F-TYPE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF018174; AAC04671.1; --  
DR PIR; T07837; T07837.  
DR HSSP; P10599; 1AIU.  
DR InterPro; IPR006662; Thioresox\_dom2.  
DR Pfam; PF00085; Thioresox\_1.  
DR PRINTS; PR00421; THIOREDOXIN.  
DR PROSITE; PS00194; THIOREDOXIN; 1.  
KW Redox-active center; Electron transport; Chloroplast; Transit peptide.  
FT CHAIN 1 ?  
FT CHAIN 2 ?  
FT CHAIN 102 105 THIOREDOXIN F-TYPE.  
FT DISULFID 102 105 REDOX-ACTIVE (BY SIMILARITY).  
SQ SEQUENCE 182 AA; 19757 MW; 9EE3E6AC7C0F35A6 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILKDN 24  
Db 149 KILKDN 154

## RESULT 103

ID THIF PEA STANDARD; PRT; 182 AA.  
AC P29450;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Thioresoxin F-type, chloroplast precursor (TRX-F).  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
[1]

RP SEQUENCE FROM N.A.  
RA TISSUE=Leaf;  
RX MEDLINE=92256804; PubMed=1581563;  
RA Lepintec L., Hodges M., Gadal P., Cretin C.;  
RT "Isolation, characterization and nucleotide sequence of a full-length  
RT pea cDNA encoding thioresoxin-f.";  
RL Plant Mol. Biol. 18:1023-1025(1992).  
[2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=96215867; PubMed=8642611;  
RA Sahravy M., Hecht V., Lopez Jaramillo J., Chueca A., Chartier Y.,  
RA Meyer Y.;  
RT "Intron position as an evolutionary marker of thioresoxins and  
RT thioresoxin domains.";  
RL J. Mol. Evol. 42:422-431(1996).

CC -!- FUNCTION: Participates in various redox reactions through the  
CC reversible oxidation of the active center dithiol to a disulfide.  
CC The F form is known to activate a number of enzymes of the  
CC photosynthetic carbon cycle.  
CC -!- SUBCELLULAR LOCATION: Chloroplast.  
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. PLANT F-TYPE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X63537; CAA45098.1; --  
DR EMBL; U35830; AAC49357.1; --  
DR PIR; S20929; S20929.  
DR HSSP; P10599; 1AIU.  
DR InterPro; IPR006662; Thioresox.  
DR InterPro; IPR006663; Thioresox\_dom2.  
DR Pfam; PF00085; Thioresox\_1.  
DR PRINTS; PR00421; THIOREDOXIN.  
DR PROSITE; PS00194; THIOREDOXIN; 1.  
KW Redox-active center; Electron transport; Chloroplast; Transit peptide.

FT CHAIN 1 69 CHLOROPLAST (POTENTIAL).  
FT TRANSIT 70 182 THIOREDOXIN F-TYPE.  
FT CHAIN 106 109 REDOX-ACTIVE (BY SIMILARITY).  
FT DISULFID 106 109 REDOX-ACTIVE (BY SIMILARITY).  
SQ SEQUENCE 182 AA; 19775 MW; 158FC352CB9E0FF1 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILKDN 24  
Db 153 KILKDN 158

```

RESULT 104
EFP_BACFR
ID EFP_BACFR STANDARD; PRT; 185 AA.
AC P70389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor P (EFP).
GN EFP
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN SEQUENCE FROM N.A.
RC STRAIN=BF1.
RX MEDLINE=98311074; PubMed=9648740;
RA Abratt V.R., Mbewe M., Woods D.R.;
RT "Cloning of an EF-P homologue from Bacteroides fragilis that
RT increases B. fragilis glutamine synthetase activity in Escherichia
RT coli."
RL Mol. Gen. Genet. 258:363-372(1998)
CC -!- FUNCTION: Involved in peptide bond synthesis. Stimulates efficient
CC translation and peptide-bond synthesis on native or reconstituted
CC 70S ribosomes in vitro. Probably functions indirectly by altering
CC the affinity of the ribosome for aminoacyl-tRNA, thus increasing
CC their reactivity as acceptors for peptidyl transferase (By
CC similarity).
CC -!- PATHWAY: Protein biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the elongation factor P family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75509; AAC26328.1; -.
CC HAMAP; MF_00141; -.
CC DR InterPro; IPR001059; EFP-P.
CC DR Pfam; PF01132; EFP; 1.
CC DR TIGRFAMs; TIGR00038; efp; 1.
CC DR PROSITE; PS01275; EFP; 1.
CC DR Protein biosynthesis, Elongation factor.
CC KW SEQUENCE 185 AA; 20812 MW; 2457150F941A240E CRC64;
CC
CC Query Match 1.8%; Score 6; DB 1; Length 185;
CC Best Local Similarity 100.0%; Pred. No. 2.2e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
QY 300 KLIEFL 305
DB 18 KLIEFL 23
|||
ID THF2_ARATH STANDARD; PRT; 185 AA.
AC Q9XFH9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioedoxin F-type 2, chloroplast precursor (TRX-F2).
GN AT5G16400 OR MQK4.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi6.
OX NCBI_TaxID=3702;

```

```

RN SEQUENCE FROM N.A.
RP Verdoucq J., Meyer Y.;
RA "Characterisation of the Arabidopsis thaliana thioedoxin f family.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT pl clones";
RL DNA Res. 4:215-230(1997).
CC -!- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of the active center dithiol to a disulfide.
CC The F form is known to activate a number of enzymes of the
CC photosynthetic carbon cycle (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THIOEDOXIN FAMILY. PLANT P-TYPE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF144386; AAD35004.1; -.
CC DR EMBL; AB005242; BAB09607.1; -.
CC DR HSSP; P10599; 1ATU.
CC DR InterPro; IPR006662; Thioed.
CC DR InterPro; IPR006663; Thioedox_dom2.
CC DR Pfam; PF00085; thioed; 1.
CC DR PRINTS; PR00421; THIOEDOXIN.
CC DR PROSITE; PS00194; THIOEDOXIN; 1.
CC KW Redox-active center; Electron transport; Chloroplast; Transit peptide;
CC Multigene family.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC FT CHAIN 1 ? 185 THIOEDOXIN F-TYPE 2.
CC FT DISULFID 109 112 REDOX-ACTIVE (BY SIMILARITY).
CC SQ SEQUENCE 185 AA; 19999 MW; DB7EA3FAA35EDD5C CRC64;
CC
CC Query Match 1.8%; Score 6; DB 1; Length 185;
CC Best Local Similarity 100.0%; Pred. No. 2.2e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
QY 19 KILKDN 24
DB 156 KILKDN 161
|||
ID VMA2_TRTV STANDARD; PRT; 186 AA.
AC P33494;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Matrix glycoprotein M2 (Envelope-associated 22 kDa protein).
GN 22K.
OS Turkey rhinotracheitis virus (TRTV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=11264;
RN SEQUENCE FROM N.A.
RP MEDLINE=9233255; PubMed=1629697;
RA Ling R., Easton A.J., Pringle C.R.;
RT "Sequence analysis of the 22K, SH and G genes of turkey
RT rhinotracheitis virus and their intergenic regions reveals a gene

```

```

order different from that of other pneumoviruses.";
RT J. Gen. Virol. 73:1709-1715(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UK/3BV/85;
RX MEDLINE=92300329; PubMed=1607858;
RA Yu Q., Davis P.J., Brown T.D.K., Cavanagh D.;
RT "Sequence and in vitro expression of the M2 gene of turkey
RT rhinotracheitis pneumovirus.";
RL J. Gen. Virol. 73:1355-1363(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S40185; AAB22544.1; -.
DR EMBL; X63408; CAA45004.1; -.
DR PIR; J01623; J01623.
DR PIR; J01987; J01987.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; Zf-CCH; 1.
DR SMART; SM00356; Znf_C3H; 1.
KW Matrix protein; Envelope protein; Glycoprotein.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 175 175 N -> S (IN REF. 2).
SQ SEQUENCE 186 AA; 20986 MW; BC638B5741E612E5 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160
DB 158 LAKIIL 163

RESULT 107
MOBA_CAMJE STANDARD; PRT; 191 AA.
AC Q9PMU9;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable molybdopterine-guanine dinucleotide biosynthesis protein A.
GN MOBA OR CUL350.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagals K., Karyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterine (MPT)
CC forming molybdopterine guanine dinucleotide (MGD) (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
RT J. Gen. Virol. 73:1709-1715(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UK/3BV/85;
RX MEDLINE=92300329; PubMed=1607858;
RA Yu Q., Davis P.J., Brown T.D.K., Cavanagh D.;
RT "Sequence and in vitro expression of the M2 gene of turkey
RT rhinotracheitis pneumovirus.";
RL J. Gen. Virol. 73:1355-1363(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S40185; AAB22544.1; -.
DR EMBL; X63408; CAA45004.1; -.
DR PIR; J01623; J01623.
DR PIR; J01987; J01987.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; Zf-CCH; 1.
DR SMART; SM00356; Znf_C3H; 1.
KW Matrix protein; Envelope protein; Glycoprotein.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 175 175 N -> S (IN REF. 2).
SQ SEQUENCE 186 AA; 20986 MW; BC638B5741E612E5 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160
DB 158 LAKIIL 163

RESULT 107
MOBA_CAMJE STANDARD; PRT; 191 AA.
AC Q9PMU9;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable molybdopterine-guanine dinucleotide biosynthesis protein A.
GN MOBA OR CUL350.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagals K., Karyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: Links a guanosine 5'-phosphate to molybdopterine (MPT)
CC forming molybdopterine guanine dinucleotide (MGD) (By similarity).
CC -!- PATHWAY: Molybdenum cofactor biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the mobA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S40185; AAB22544.1; -.
DR EMBL; X63408; CAA45004.1; -.
DR PIR; J01623; J01623.
DR PIR; J01987; J01987.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; Zf-CCH; 1.
DR SMART; SM00356; Znf_C3H; 1.
KW Matrix protein; Envelope protein; Glycoprotein.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 175 175 N -> S (IN REF. 2).
SQ SEQUENCE 186 AA; 20986 MW; BC638B5741E612E5 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 FLEQNY 206
DB 112 FLEQNY 117

RESULT 108
Y4D2_METAC STANDARD; PRT; 198 AA.
ID Y4D2_METAC
AC Q8TIV9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MA4032.
GN MA4032.
OS Methanococcus acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Linton N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Zimmer A., Barber R.D., Cann I., Graham D.B., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -!- SIMILARITY: Belongs to the UPF0228 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE011115; AAM07380.1; -.
DR InterPro; IPR008887; UPF0228.
DR Pfam; PF05727; UPF0228; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 198 AA; 23463 MW; 9D5B2C3A07DC6918 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LAILEK 30

```

```
Db 139 LAILEK 144
|||||
RESULT 109
AA27 HUMAN
ID AA27 HUMAN STANDARD; PRT; 199 AA.
AC O60232;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sjogren's syndrome/scleroderma autoantigen 1 (Autoantigen p27).
GN SSSCAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98145811; PubMed=9486406;
RA Muro Y., Yamada T., Himeno M., Sugimoto K.;
RT "cDNA cloning of a novel autoantigen targeted by a minor subset of
RT anti-centromere antibodies.";
RL Clin. Exp. Immunol. 111:372-376 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bask S.A., McQuellan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Might play a role in mitosis. Antigenic molecule. Could
CC be a centromere-associated protein. May induce anti-centromere
CC antibodies.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001740; BAA25263.1; -
CC
CC DR EMBL; BC014791; AAH14791.1; -
CC
CC DR Genew; HGNC:11328; SSSCAL.
CC
CC DR GK; O60232; -
CC
CC DR MIM; 606044; -
CC
CC DR GO; GO:0007067; P:mitosis; TAS.
CC
CC Antigen; Mitosis.
CC
CC KW DOMAIN 75 ASP/GLU-RICH (ACIDIC).
CC
CC FT DOMAIN 113 122 PRO-RICH.
CC
CC FT DOMAIN 138 145 PRO-RICH.
CC
CC FT DOMAIN 170 177 SER/THR-RICH.
CC
CC SEQUENCE 199 AA; 21474 MW; EF08439FDFFDALDAB CRC64;

Query Match 1.8%; Score 6; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 EPPTFA 65
Db 14 EPPTFA 19

RESULT 110
AA27 MOUSE
ID AA27 MOUSE STANDARD; PRT; 199 AA.
AC P56873;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sjogren's syndrome/scleroderma autoantigen 1 homolog (Autoantigen p27
DE homolog).
GN C184L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ICB; TISSUE=Forebrain;
RX MEDLINE=99443733; PubMed=10512749;
RA Sakuma-Takagi M., Tohyama Y., Kasama-Yoshida H., Sakagami H.,
RA Kondo H., Kurihara T.;
RT "Novel related cDNAs (C184L, C184M, and C184S) from developing mouse
RT brain encoding two apparently unrelated proteins.";
RL Biochem. Biophys. Res. Commun. 263:737-742 (1999).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE EARLY POSTNATAL BRAIN.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE FOREBRAIN AT E16.
CC -!- MISCELLANEOUS: TWO ORF ARE CONTAINED IN THE C184L GENE, BUT ONLY
CC THIS PROTEIN IS EFFICIENTLY TRANSLATED AND EXPRESSED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB012727; BAA87048.1; -
CC
CC DR SEQUENCE 199 AA; 21336 MW; 13A5222048009E00 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 EPPTFA 65
Db 14 EPPTFA 19

RESULT 111
HAMI CAMJE
ID HAMI CAMJE STANDARD; PRT; 200 AA.
AC Q9PMS6; Q9ZRF65;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE HAMI protein homolog.
GN CUI374C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OH4384;

Query Match 1.8%; Score 6; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 EPPTFA 65
Db 14 EPPTFA 19
```

RA Gilbert M., Michniewicz J., Wakarchuk W.W.;  
RT "Cloning of a multidrug-efflux transporter homolog from *Campylobacter*  
RL *jejuni*";  
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagels K., Karyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Whitehead A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whittam S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*  
RL Nature 403:665-668(2000).  
CC -!- SIMILARITY: Belongs to the HAM1 NTPase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF113952; AAD10059.1; -;  
CC EMBL; AL139078; CAB73801.1; -;  
CC PIR; C81282; C81282.  
CC HSSP; Q57679; 1B78.  
CC HAMAP; MF\_014005; -; 1.  
CC InterPro; IPR002637; Hamplp like.  
CC Pfam; PF01725; Hamplp like; 1.  
CC ProDom; PD004952; Hamplp like; 1.  
CC DR TIGRFAMs; TIGR00042; TIGR00042; 1.  
CC KW Hydrolyase; Complete proteome.  
CC FT CONFLICT 120 120 Y -> H (IN REF. 1).  
CC SQ SEQUENCE 200 AA; 22375 MW; 02F29AF6FEB0FBA1 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 155 LAKIIL 160  
Db 189 LAKIIL 194  
RESULT 112  
ID HIS5\_METKA STANDARD; PRT; 200 AA.  
AC Q8TV83;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Imidazole glycerol phosphate synthase subunit hisH (EC 2.4.2.-) (IGP  
DE synthase glutamine amidotransferase subunit) (IGP synthase subunit  
DE hisH) (ImGP synthase subunit hisH) (IGPS subunit hisH).  
GN HSH OR MK1510.  
OS Methanopyrus kandleri.  
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
OC Methanopyrus.  
OX NCBI\_TaxID=2320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE=21927647; PubMed=11930014;  
RA Slesarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,  
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
RA Natarale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.;  
RT "The complete genome of hyperthermophile *Methanopyrus kandleri* AV19  
RT and monophyle of archaeal methanogens.";

Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
-!- FUNCTION: IGP catalyzes the conversion of PRFAR and glutamine to  
IGP, AICAR and glutamate. The hisH subunit provides the glutamine  
amidotransferase activity that produces the ammonia necessary to  
hisP for the synthesis of IGP and AICAR (By similarity).  
-!- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulose-1-  
ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-  
carboxamide + L-glutamine = imidazole-glycerol phosphate + 5-  
aminoimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.  
-!- PATHWAY: Histidine biosynthesis; fifth step.  
-!- SUBUNIT: Heterodimer of hisH and hisP (By similarity).  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE010441; AM02723.1; -;  
CC HAMAP; MF\_00278; -; 1.  
CC InterPro; IPR000991; GATase\_1.  
CC Pfam; PF00117; GATase\_1.  
CC PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
CC KW Histidine biosynthesis; Transferase; Glutamine amidotransferase;  
CC Complete proteome.  
CC FT ACT SITE 79 79 BY SIMILARITY.  
CC ACT SITE 179 179 BY SIMILARITY.  
CC ACT SITE 181 181 BY SIMILARITY.  
CC SQ SEQUENCE 200 AA; 22013 MW; E46FEAD5E28EA035 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 229 LKLLGE 234  
Db 188 LKLLGE 193  
RESULT 113  
ID YW3\_YEAST STANDARD; PRT; 201 AA.  
AC Q04772;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 23.2 kDa protein in ABF2-CHL12 intergenic region.  
GN YMR073C OR YW916.12C;  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AB972;  
RX MEDLINE=97313268; PubMed=9169872;  
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,  
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;  
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome  
RT XIII.";  
RL Nature 387:90-93(1997).  
CC -!- SIMILARITY: Belongs to the cytochrome b5 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way



CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; 248952; CAA88798.1; -.  
DR PIR; S52833; S52833.  
DR HSSP; P00175; 1FCB.  
DR GernOnline; 142739; -.  
DR SGD; S0004677; YMR073C.  
DR InterPro; IPR001199; Cyt\_B5.  
DR Pfam; PF00173; heme\_1; 1.  
DR ProDom; PD00612; Cyt\_B5; 1.  
DR PROSITE; PS00191; CYTOCHROME B5\_1; 1.  
DR PROSITE; PS0255; CYTOCHROME B5\_2; 1.  
KW Hypothetical protein; Heme.  
FT METAL 158 158 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
FT METAL 182 182 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
SQ SEQUENCE 201 AA; 23182 MW; 7686E1DEF08DDDFE CRC64;

Query Match 1.8%; Score 6; DB 1; Length 201;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0;

QY 213 YEKLQ 218  
DQ 187 YEKLQ 192

-----  
RESULT 114  
TRPF\_BACCR STANDARD; PRT; 202 AA.  
AC Q81GG6;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE N-(5'-phosphoribosyl)anthranilate isomerase (EC 5.3.1.24) (PRA1).  
GN TRPF OR BC1236.  
OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=226900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608415; PubMed=12721630;  
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
RA Kapatal V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,  
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas I.,  
RA Grachkin Y., Pusch G., Haselkorn R., Forstner S.D.,  
RA Overbeek R., Kyrpides N.;  
RT "Genome sequence of Bacillus cereus and comparative analysis with  
Bacillus anthracis.";  
RL Nature 423:87-91(2003).  
CC -!- CATALYTIC ACTIVITY: N-(5-phospho-beta-D-riboseyl)-anthranilate = 1-  
(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate.  
CC -!- PATHWAY: Tryptophan biosynthesis; third step.  
CC -!- SIMILARITY: Belongs to the trpF family.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; AF017001; AAP08221.1; -.  
DR HANAP; MF\_00135; -; 1.  
DR InterPro; IPR001240; PRA1.  
DR Pfam; PF00697; PRA1; 1.  
KW Isomerase; Tryptophan biosynthesis; Complete proteome.  
SQ SEQUENCE 202 AA; 22556 MW; 42EB66CEB3513048 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 202;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 92 EGKQDV 97  
DQ 182 EGKQDV 187

-----  
RESULT 115  
TFAB\_ECOLI STANDARD; PRT; 203 AA.  
AC Q47427;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Tail fiber assembly protein homolog.  
GN T.  
OS Escherichia coli.  
OG Plasmid p15B.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=15T-;  
RX MEDLINE=92051368; PubMed=1945872;  
RA Sandmeier H., Iida S., Huebner P., Hiestadt-Nauer R., Arber W.;  
RT "Gene organization in the multiple DNA inversion region min of plasmid  
p15B of E.coli 15T-: assemblage of a variable gene.";  
RL Nucleic Acids Res. 19:5831-5838(1991).  
CC -!- SIMILARITY: BELONGS TO THE TFA FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; X62121; CAA44047.1; -.  
DR PIR; S18684; S18684.  
DR InterPro; IPR003458; DUF144.  
DR Pfam; PF02413; Caudo\_TAP; 1.  
KW Plasmid.  
SQ SEQUENCE 203 AA; 22198 MW; 7D7245C2BE2403DC CRC64;

Query Match 1.8%; Score 6; DB 1; Length 203;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0;

QY 224 TKRQSL 229  
DQ 137 TKRQSL 142

-----  
RESULT 116  
YEV8\_YEAST STANDARD; PRT; 203 AA.  
AC P40080;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 23.5 kDa protein in LCPS-PAK1 intergenic region.  
GN YER128W OR YSGP-ORF44.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288C / AB972;  
RX MEDLINE=97313264; PubMed=9169868;  
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,

RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
RA Lin D., Masedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.,  
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V";  
RL Nature 387:78-81(1997).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U18916; AAC03226.1; --  
CC PIR; S43222; S43222; --  
CC GERMOnline; 139207; --  
CC SGD; S0000930; YER128W.  
KW Hypothetical protein.  
SQ SEQUENCE 203 AA; 23488 MW; 819930C5DD56B1D CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 203;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0;  
QY 64 EVAQQL 69  
Db 62 EVAQQL 67  
|||||  
SEQUENCE FROM N.A.  
RX MEDLINE=22608414; PubMed=12721629;  
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,  
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilestone J., Wu M.,  
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,  
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,  
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Nierman J.F.,  
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,  
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,  
RA Fraser C.M.;  
RT The genome sequence of *Bacillus anthracis* Ames and comparison to  
RT closely related bacteria.";  
RL Nature 423:81-86(2003).  
CC -!- CATALYTIC ACTIVITY: N-(5-phospho-beta-D-ribose)-L-anthranilate = 1-  
CC (2-carboxyphenylamino)-1-deoxy-D-ribose 5-phosphate.  
CC -!- PATHWAY: tryptophan biosynthesis; third step.  
CC -!- SIMILARITY: Belongs to the trpF family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; AE017028; AAP25210.1; --  
CC TIGR; BAI252; --  
CC HAMAP; MF 00135; --; 1.  
CC InterPro; IPR001240; PRAI.  
CC Pfam; PF00697; PRAI; 1.  
KW Isomerase; Tryptophan biosynthesis; Complete proteome.  
SQ SEQUENCE 204 AA; 22619 MW; F153422997A4D765 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 204;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0;  
QY 92 EGKXDV 97  
Db 184 EGKXDV 189  
|||||  
SEQUENCE FROM N.A.  
RX MEDLINE=96163468; PubMed=8575443;  
RA van Belzen N., Diesveld M.P.G., van der Made A.C.J., Nozawa Y.,  
RA Dirjens W.N.M., Vlietstra R., Trapman J., Bosman F.T.;  
RT Identification of mRNAs that show modulated expression during colon  
RT carcinoma cell differentiation.";  
RL Eur. J. Biochem. 234:843-848(1995).  
[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Lymph.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blacklesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- TISSUE SPECIFICITY: Down-regulated during the in vitro  
CC differentiation of HT29-D4 colon carcinoma cells.  
CC -!- SIMILARITY: TO C.ELEGANS R02F2.2 AND YEAST YOL114C.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X81788; CA857387.1; -.  
DR EMBL: BC015335; AAH15335.1; -.  
DR PIR: S63540; S63540.  
DR Genew: HGNC:5359; ICT1.  
DR MIM: 603000; -.  
DR InterPro: IPR000352; Pep\_rel\_factor\_I.  
DR Pfam: PF00472; RF-1; 1.  
SQ SEQUENCE 206 AA; 23630 MW; 663BF52443D41540 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 LGELIL 237  
|||||  
Db 130 LGELIL 135

## RESULT 119

ID PCP2\_THETN STANDARD; PRT; 206 AA.  
AC Q8R3J6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Pyrolydione-carboxylate peptidase 2 (EC 3.4.19.3) (5-oxoprol-yl-peptidase 2) (Pyroglutamyl-peptidase 1 2) (PGP-I 2) (Pyrase 2).  
GN PCP2 OR TIE1612.  
OS Thermoanaerobacter tengcongensis.  
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4 / JCM 11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
Tan H., Chen R., Wang J., Yu J., Yang H.;  
"A complete sequence of T. tengcongensis genome.";  
RL Genome Res. 12:689-700(2002).  
CC -!- FUNCTION: Removes 5-oxoproline from various penultimate amino acid residues except L-proline (By similarity).  
CC -!- CATALYTIC ACTIVITY: 5-oxoprol-yl-peptide + H(2)O = 5-oxoprol-yl-peptide +  
peptide.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family C15.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE013116; AM24815.1; -.  
DR HAMAP: MF\_00417; -.  
DR InterPro: IPR000816; Peptidase C15.  
DR Pfam: PF01470; Peptidase C15; 1.  
DR ProDom: PD008480; Peptidase C15; 1.  
DR TIGRfam: TIGR00504; PYRO\_pbase; 1.  
DR PROSITE: PS01334; PYRASE\_C15; 1.  
DR PROSITE: PS01333; PYRASE\_GLU; 1.  
KW Hydrolase; Thiol protease; Complete proteome.  
FT SITE 78 BY SIMILARITY.  
ACT SITE 141 141 BY SIMILARITY.  
FT ACT\_SITE 165 165 BY SIMILARITY.

SQ SEQUENCE 206 AA; 22617 MW; 3D9F4E9DFDB34DAD CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKI 20  
|||||  
Db 33 AEIVKI 38

## RESULT 120

ID YIOR\_CVBF STANDARD; PRT; 207 AA.  
AC P22654;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein in nucleocapsid ORF (IORF).  
OS Bovine coronavirus (strain F15) (BCoV) (BCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus.  
OX NCBI\_TaxID=11129;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89087718; PubMed=3207501;  
RA Cruciere C., Laporte J.;  
"Sequence and analysis of bovine enteric coronavirus (F15) genome.  
I. Sequence of the gene coding for the nucleocapsid protein; analysis of the predicted protein.";  
RL Ann. Inst. Pasteur Virol. 139:123-138(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

SQ SEQUENCE 207 AA; 23001 MW; A4E5DE61171B4B50 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 207;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDK 265  
|||||  
Db 59 NLLRDK 64

## RESULT 121

ID IF2B\_THEAC STANDARD; PRT; 209 AA.  
AC Q9HKJ3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Probable translation initiation factor 2 beta subunit (eIF-2-beta).  
GN EIF2B OR TA605.  
OS Thermoplasma acidophilum.  
OC Archaea; Euryarchaeota; Thermoplasma; Thermoplasmatales;  
OC Thermoplasmataceae; Thermoplasma.  
OX NCBI\_TaxID=2303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 1728;  
RX MEDLINE=20479972; PubMed=11029001;

RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.,  
RT "The genome sequence of the thermophilic scavenger Thermoplasma  
acidophilum".  
RL Nature 407:508-513(2000).  
CC -!- FUNCTION: eIF-2 functions in the early steps of protein synthesis  
CC by forming a ternary complex with GTP and initiator tRNA (By  
CC similarity).  
CC -!- SUBUNIT: Heterotrimer composed of an alpha, a beta and a gamma  
CC chain (By similarity).  
CC -!- SIMILARITY: Belongs to the eIF-2-beta / eIF-5 family.  
CC -!- SIMILARITY: Contains 1 TRAM domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AL445064; CAC11744.1; --  
DR HAMAP; MF\_00232; fused; 1.  
DR InterPro; IPR002735; eIF5\_eIF2B.  
DR InterPro; IPR004458; TIF\_aIF-2beta.  
DR InterPro; IPR002792; TRAM.  
DR Pfam; PF01873; eIF5\_eIF2B; 1.  
DR Pfam; PF01938; TRAM; 1.  
DR ProDom; PD004078; eIF5\_eIF2B; 1.  
DR SMART; SM00653; eIF2B\_5; 1.  
DR TIGRFAMs; TIGR00311; aIF-2beta; 1.  
DR PROSITE; PS0926; TRAM; 1.  
KW Initiation factor; Protein biosynthesis; Complete proteome.  
FT DOMAIN 144 202 TRAM.  
SQ SEQUENCE 209 AA; 23899 MW; 577234B9BEC61B8D CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 209;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 212 DYKLL 217  
DB 4 DYKLL 9  
  
RESULT 122  
GTP2\_BUFBU STANDARD; PRT; 210 AA.  
AC P83325;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glutathione S-transferase P 2 (EC 2.5.1.18) (BGGSTP2-2) (GST class-  
DE pi).  
OS Bufo bufo (European toad).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Bufonidae; Bufo.  
OX NCBI\_TaxID=8384;  
RN [1]  
RP SEQUENCE, FUNCTION, CATALYTIC ACTIVITY, SUBUNIT, AND TISSUE  
RP SPECIFICITY.  
RC TISSUE=Liver;  
RX MEDLINE=22123471; PubMed=12127579;  
RA Bucciarelli T., Sacchetta P., Amicarelli F., Petruzzelli R.,  
RA Melino S., Rotilio D., Celli N., Di Lio C.;  
RT "Amino acid sequence of the major form of toad liver glutathione  
RT transferase".  
RL Int. J. Biochem. Cell Biol. 34:1286-1290(2002).  
CC -!- FUNCTION: Conjugation of reduced glutathione to a wide number  
CC of exogenous and endogenous hydrophobic electrophiles.  
CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.  
CC -!- SUBUNIT: Homodimer.  
CC -!- TISSUE SPECIFICITY: Liver, kidney, muscle, skin, lung and

CC ovary.  
CC -!- SIMILARITY: Belongs to the GST superfamily. Pi family.  
DR InterPro; IPR004046; GST\_Cterm.  
DR InterPro; IPR004045; GST\_Nterm.  
DR InterPro; IPR003082; GST\_pi.  
DR Pfam; PF00043; GST\_C; 1.  
DR Pfam; PF02798; GST\_N; 1.  
DR PRINTS; PR01268; GSTNSFRASEP.  
DR Transferrase; Multigene family.  
SQ SEQUENCE 210 AA; 24178 MW; 3D233F9AE53F3F8C CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 210;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 329 IRDLKK 334  
DB 42 IRDLKK 47  
  
RESULT 123  
HIS1\_SINY3 STANDARD; PRT; 210 AA.  
ID HIS1\_SINY3  
AC Q55503;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP phosphoribosyltransferase (EC 2.4.2.17).  
DS HISG OR SLL0900.  
GN Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugitara M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 64% to 92% of the genome."  
RL DNA Res. 2:153-166(1995).  
CC -!- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribose)-ATP + diphosphate =  
CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.  
CC -!- PATHWAY: Histidine biosynthesis; first step. Very important in the  
CC regulation of histidine metabolism.  
CC -!- SUBUNIT: Homohexamer (By similarity). Binds to hisZ possibly to  
CC allow the regulation of hisG transferase activity by histidine (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- DOMAIN: Lacks the C-terminal regulatory region which is replaced  
CC by hisZ.  
CC -!- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.  
CC Short subfamily.  
  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: D64006; BAA10855.1; --  
DR PIR; S76008;  
DR HAMAP; MF\_01018; --; 1.  
DR InterPro; IPR001348; ATP\_phospho\_trans.  
DR Pfam; PF01634; HisG; 1.  
DR ProDom; PD003516; ATP\_phospho\_trans; 1.  
DR TIGRFAMs; TIGR00070; hisG\_1-  
DR PROSITE; PS01316; ATP\_P\_PHORIBOSYLTR; 1.  
KW Histidine biosynthesis; Transferase; Glycosyltransferase;  
KW Complete proteome.  
SQ SEQUENCE 210 AA; 23436 MW; C8CEBD1EC8C1ABF0 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 210;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 AVAQLA 70  
|||||  
Db 81 AVAQLA 86

RESULT 124  
CCGX\_MOUSE STANDARD; PRT; 211 AA.  
AC Q9JUV3; 2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Voltage-dependent calcium channel gamma-like subunit (Neuronal  
DE voltage-gated calcium channel gamma-like subunit).  
GN P1 OR PR OR CACNG5.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Kidney; PubMed=10734232;  
RX MEDLINE=20200313; PubMed=10734232;  
RA Klugbauer N., Dai S., Specht V., Lacinova L., Marais E., Bohn G.,  
RA Hofmann F.;  
RT "A family of gamma-like calcium channel subunits.";  
RL FEBS Lett. 470:189-197(2000).  
CC -!- FUNCTION: Thought to stabilize the calcium channel in an  
CC inactivated (closed) state (by similarity).  
CC -!- SUBUNIT: The L-type calcium channel is composed of five subunits:  
CC alpha-1, alpha-2/delta, beta and gamma.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -!- SIMILARITY: Belongs to the PMP-22 / EMP / MP20 family. CACNG  
CC subfamily.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AJ272046; CAB86387.1; -.  
DR MGD; MGI:2157899; Pr1.  
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;  
KW Calcium channel.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 96 116 POTENTIAL.  
FT TRANSMEM 131 151 POTENTIAL.  
FT TRANSMEM 155 175 POTENTIAL.  
SQ SEQUENCE 211 AA; 23230 MW; ALA05270B971CC03 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 211;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 ILLKNO 298  
|||||  
Db 149 ILLKNO 154

RESULT 125  
MOTA\_BPT4 STANDARD; PRT; 211 AA.  
ID MOTA\_BPT4  
AC M22915;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Transcription regulatory protein mota (Middle transcription protein  
DE A).  
GN MOTA.  
OS Bacteriophage T4.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=10665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91141300; PubMed=2287273;  
RA Uzan M., Brody E., Favre R.;  
RT "Nucleotide sequence and control of transcription of the  
RT bacteriophage T4 mota regulatory gene.";  
RL Mol. Microbiol. 4:1487-1496(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22514363; PubMed=12626685;  
RA Miller E.S., Kutter E., Mosig G., Ariesaka F., Kunisawa T., Ruger W.;  
RT "Bacteriophage T4 genome.";  
RL Microbiol. Mol. Biol. Rev. 67:86-156(2003).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.19 ANGSTROMS) OF 2-96.  
RX MEDLINE=97299876; PubMed=9155025;  
RA Finnin M.S., Cicero M.P., Davies C., Porter S.J., White S.W.,  
RA Kreuzer K.N.;  
RT "The activation domain of the Mota transcription factor from  
RT bacteriophage T4.";  
RL EMBO J. 16:1992-2003(1997).  
CC -!- FUNCTION: Required for the transcriptional activation of  
CC middle promoters. Middle promoters are characterized by the  
CC presence of the conserved sequence [AT]3TGCCTNA (Mota box).  
CC Mota binds directly to Mota boxes.  
CC -!- MISCELLANEOUS: Mota synthesis starts immediately after infection  
CC and stops abruptly some 4-5 min later.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; Z48569; CAA88453.1; -.  
DR EMBL; AF158101; AAD42607.1; -.  
DR PIR; JVO101; ZKEPT4.  
DR PDB; 1BUA; 18-NOV-98.  
DR PDB; 1IIS; 18-APR-01.  
DR PDB; 1KAF; 24-APR-02.  
KW Transcription regulation; Activator; DNA-binding; 3D-structure.  
FT DNA BIND 23 42 H-T-H MOTIF (BY SIMILARITY).  
FT MUTAGEN 140 140 A->D: TEMPERATURE-SENSITIVE.  
FT HELIX 4 10  
FT TURN 11 13  
FT TURN 17 28  
FT TURN 30 31  
FT STRAND 32 32  
FT HELIX 34 39  
FT TURN 40 40  
FT TURN 42 43  
FT HELIX 46 57  
FT TURN 58 60  
FT STRAND 62 65  
FT TURN 66 67  
FT STRAND 68 71  
FT STRAND 73 89  
FT TURN 90 90  
FT HELIX 91 94  
SQ SEQUENCE 211 AA; 23574 MW; 841625941A194E3F CRC64;

Query Match 1.8%; Score 6; DB 1; Length 211;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LAILEK 30
Db 139 LAILEK 144

RESULT 126
PUR3_HAEN STANDARD; PRT; 212 AA.
AC P43846;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2) (GART) (GAR
DE transforlylase) (5'-phosphoribosylglycinamide transferase)
GN PURN OR H11428.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + N(1)-(5-phospho-D-
CC ribosyl)glycinamide = tetrahydrofolate + N(2)-formyl-N(1)-(5-
CC phospho-D-ribose)glycinamide.
CC -!- PATHWAY: De novo purine biosynthesis; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: TO OTHER GART FROM BACTERIA AND EUKARYOTES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U32822; AAC23075.1; -.
CC DR PIR; F64122; F64122.
CC DR HSSP; P08179; LGAR.
CC DR TIGR; H11428; -.
CC
CC InterPro; IPR002376; formyl transf.
CC InterPro; IPR001555; GART_AS.
CC InterPro; IPR004607; FurN.
CC Pfam; PF00551; formyl transf. 1.
CC TIGRFAMs; TIGR00639; FurN; 1.
CC PROSITE; PS00373; GART; 1.
CC Purine biosynthesis; Transferase; Complete proteome.
FT ACT_SITE 145 145 BY SIMILARITY.
SQ SEQUENCE 212 AA; 23463 MW; D1AE249CDD8B79C4 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LKDNLA 26
Db 192 LKDNLA 197
```

```
RESULT 127
RADB_METJA STANDARD; PRT; 212 AA.
AC Q57702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA repair and recombination protein radB.
GN RADB OR MJ0254.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii."
CC Science 273:1058-1073(1996).
CC -!- FUNCTION: Involved in DNA repair and in homologous recombination.
CC May regulate the cleavage reactions of the branch-structured DNA.
CC Has a very weak ATPase activity that is not stimulated by DNA.
CC Binds DNA but does not promote DNA strands exchange (By
CC similarity).
CC -!- SIMILARITY: Belongs to the eukaryotic recA-like protein family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U67480; AAB98241.1; -.
CC DR PIR; G64331; G64331.
CC DR TIGR; MJ0254; -.
CC DR HAMAP; MF_00350; -.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR001553; RecA.
CC DR SMART; SM00382; AAA; 1.
CC DR PROSITE; PS0162; RECA_2; 1.
CC DNA damage; DNA recombination; ATP-binding; DNA-binding;
CC Complete proteome.
FT NP_BIND 20 27 ATP (POTENTIAL).
SQ SEQUENCE 212 AA; 23842 MW; 0BD1C9D5D68579CB CRC64;

Query Match 1.8%; Score 6; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LAILEK 30
Db 185 LAILEK 190

RESULT 128
MMCI_METAC STANDARD; PRT; 217 AA.
ID MMCI_METAC
AC P58867;
DT 28-FEB-2003 (Rel. 41, Created)
```

28-FEB-2003 (Rel. 41, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Monomethylamine corrinoid protein 1 (MWCP 1).  
MTWC1 OR MA0145.  
OS Methanosarcina acetivorans.  
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2214;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE=21929760; PubMed=11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,  
RA Linton L., McEwan P., McKernan K., Talanas J., Tirrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cunniff I., Graham D.E., Guss A.M.,  
RA Hederich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.W., Birren B.;  
RA "The Genome of Methanosarcina acetivorans reveals extensive metabolic  
RT and physiological diversity.";  
RL Genome Res. 12:532-542(2002).  
CC  
CC -1- FUNCTION: Acts as a methyl group carrier between mtmb and mtba (By  
CC similarity).  
CC -1- PATHWAY: Methanogenesis from monomethylamines.  
CC -1- SUBUNIT: Can form a complex with mtmb (By similarity).  
CC -1- SIMILARITY: Belongs to the methylamine corrinoid protein family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE010671; AM03598.1; -;  
DR InterPro; IPR006158; B12-binding.  
DR InterPro; IPR003759; CoMet synth\_B12.  
DR Pfam; PF02310; B12-binding; 1.  
DR Pfam; PF02607; B12-binding; 2; 1.  
DR Cobalt; Methanogenesis; Complete proteome.  
FT INIT.MET 0 BY SIMILARITY.  
FT  
SQ SEQUENCE 217 AA; 23016 MW; A4266EE52DD53E5 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 217;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 67 AQLAQE 72  
Db 22 AQLAQE 27  
RESULT 129  
YB63\_METUA  
ID\_YB63\_METUA STANDARD; PRT; 217 AA.  
AC Q58563;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical metal-dependent hydrolase MJ1163.  
GN MJ1163.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.

STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
MEDLINE=96337999; PubMed=8688087;  
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Bult C.J., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Karlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";  
RL Science 273:1058-1073(1996).  
CC  
CC -1- SIMILARITY: Belongs to the UPF0173 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U67558; XAB99165.1; -;  
DR PIR; B64445; B64445.  
DR TIGR; MJ1163; -;  
DR HAMAP; MF\_00457; -; 1.  
DR InterPro; IPR001279; Blactmase-like.  
DR Pfam; PF00753; lactamase\_B; 1.  
KW Hypothetical protein; Hydrolase; Complete proteome.  
SQ SEQUENCE 217 AA; 23516 MW; A2FE84E38D9BC616 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 217;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 135 APQIAL 140  
Db 146 APQIAL 151  
RESULT 130  
FTSE\_HAEIN  
ID\_FTSE\_HAEIN STANDARD; PRT; 218 AA.  
AC P44871;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cell division ATP-binding protein ftse.  
GN FTSE OR HI0769.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rd / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: NOT KNOWN. IS CODED IN AN OPERON ESSENTIAL FOR  
CC CELL DIVISION (BY SIMILARITY).  
CC -1- SIMILARITY: Belongs to the ABC transporter family.



```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32760; AAC22427.1; -.
DR PIR; H64091; H64091.
DR TIGR; H10769; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005286; IISF.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00960; 3a0501s02; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Cell division; ATP-binding; Complete proteome.
FT NP_BIND 36 43
FT NP_BIND 43 43
SQ SEQUENCE 218 AA; 24349 MW; 6E3DFE39C03F33EB CRC64;

Query Match 1.8%; Score 6; DB 1; Length 218;
Best Local Similarity 100.0%; Pred.No. 2.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110
DB 78 LRRQIG 83

RESULT 131
ENGB VIBCH
ID ENGB VIBCH STANDARD; PRT; 220 AA.
AC Q9KVN0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable GTP-binding protein engb.
GN ENGB OR VC0111.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickley E.K., Peterson J.D., Unwayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolova M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Necessary for normal cell division and for the
CC maintenance of normal septation (By similarity).
CC -!- SIMILARITY: Belongs to the engb family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AE004101; AAF93289.1; -.
DR PIR; E82364; E82364.
DR TIGR; VC0111; -.
DR HAMAP; MF 00321; -.
DR InterPro; IPR005289; GTP-binding_dom.
DR TIGRFAMs; TIGR00650; MG442; 1.
KW Cell division; Septation; GTP-binding; Complete proteome.
FT NP_BIND 34 41
FT NP_BIND 41 41
FT NP_BIND 79 83
FT NP_BIND 146 149
FT NP_BIND 149 149
SQ SEQUENCE 220 AA; 24343 MW; D4926E9E5073E37C CRC64;

Query Match 1.8%; Score 6; DB 1; Length 220;
Best Local Similarity 100.0%; Pred.No. 2.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 KRQSLK 230
DB 106 KRQSLK 111

RESULT 132
KTHY LACPL
ID KTHY LACPL STANDARD; PRT; 221 AA.
AC Q88YP6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN TMK OR LP 0703.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC -!- FUNCTION: Phosphorylation of dTMP to form dTDP in both de novo and
CC salvage pathways of dTTP synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + thymidine 5'-phosphate = ADP + thymidine
CC 5'-diphosphate.
CC -!- SIMILARITY: Belongs to the thymidylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL935253; CAD63305.1; -.
DR HAMAP; MF 00165; -.
DR InterPro; IPR000062; Thymidylate kin.
DR Pfam; PF02223; Thymidylate kin; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; 1.
KW Transferase; Kinase; Nucleotide biosynthesis; ATP-binding;
KW Complete proteome.
FT NP_BIND 11 18
FT NP_BIND 18 18
SQ SEQUENCE 221 AA; 24314 MW; C1835157FD252A3 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 221;
Best Local Similarity 100.0%; Pred.No. 2.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 236 ILDRHN 241  
Db 55 ILDRHN 60

RESULT 133  
ID FTSE\_ECOLI STANDARD; PRT; 222 AA.  
AC P10115;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cell division ATP-binding protein ftse.  
GN FTSE OR B3463 OR C4256 OR Z4837 OR ECS4312 OR SF3481 OR S4282.  
OS Escherichia coli, O6,  
OS Escherichia coli O157:H7, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334, 623;  
RN [1]  
SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12;  
RX MEDLINE=87089083; PubMed=3025556;  
RA Gill D.R., Hatfull G.F., Salmond G.P.C.;  
RT "A new cell division operon in Escherichia coli.";  
RL Mol. Gen. Genet. 205:134-145 (1986).  
RN [2]  
SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
RX MEDLINE=94316500; PubMed=8041620;  
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the  
RT region from 76.0 to 81.5 minutes."  
RL Nucleic Acids Res. 22:2576-2586 (1994).  
RN [3]  
SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
RT of uropathogenic Escherichia coli."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
RN [4]  
SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Liou S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."  
RL Nature 409:529-533 (2001).  
RN [5]  
SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12."  
RL DNA Res. 8:11-22 (2001).  
RN [6]  
SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;

RX MEDLINE=22272406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441 (2002).  
RN [7]  
SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella  
RT flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786 (2003).  
CC -!- FUNCTION: NOT KNOWN. IS CODED IN AN OPERON ESSENTIAL FOR  
CC CELL DIVISION.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X04398; CAA27985.1; -;  
CC EMBL; U00039; AAB18438.1; -;  
CC EMBL; AE000422; AAC76488.1; -;  
CC EMBL; AE016768; AAN82692.1; -;  
CC EMBL; AE005569; AAG58572.1; -;  
CC EMBL; AP002565; BAB37735.1; -;  
CC EMBL; AE015356; AAN44940.1; ALT\_INIT.  
CC EMBL; AE016992; AAP19242.1; -;  
CC PIR; H86013; H86013.  
CC PIR; H91167; H91167.  
CC PIR; S03131; CEECFE.  
CC EcoGene; EGI0340; ftse.  
CC InterPro; IPR003593; AAA ATPase.  
CC InterPro; IPR003439; ABC transporter.  
CC InterPro; IPR005286; IISF.  
CC Pfam; PF00005; ABC tran; 1.  
CC ProDom; PD000006; ABC transporter; 1.  
CC SMART; SM00382; AAA; 1.  
CC TIGRFAMs; TIGR00960; aa0501802; 1.  
CC PROSITE; PS00211; ABC TRANSPORTER 1; 1.  
CC PROSITE; PS00893; ABC TRANSPORTER 2; 1.  
CC Cell division; ATP-binding; Complete proteome.  
CC NP BIND 35 42 ATP (BY SIMILARITY).  
SQ SEQUENCE 222 AA; 24439 MW; 13CFCD8D8FE7590 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 105 LRRQIG 110  
Db 77 LRRQIG 82  
RESULT 134  
GLI4\_ARATH  
ID GLI4\_ARATH STANDARD; PRT; 222 AA.  
AC Q9M8X4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

Germin-like protein subfamily 1 member 4 precursor.  
 AT3G04180 OR T6K12.20.  
 Arabidopsis thaliana (Mouse-ear cross).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 euroids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=cv. Columbia;  
 MEDLINE=21016720; PubMed=11130713;  
 Salanoubat M., Lemcke K., Rieger M., Perez-Alonso M., Obermaier B.,  
 Fartmann B., Valle G., Bloeker H., Maché R., Puigdomenech P.,  
 Delany M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,  
 De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,  
 Wincker P., Cattolico L., Weissbach J., Saurin W., Quétier F.,  
 Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 Wurmback E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 Vezi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,  
 Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,  
 Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Clement J.,  
 Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,  
 de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
 Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
 Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
 Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,  
 Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,  
 Sasaoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
 Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana";  
 Nature 408:820-822(2000).  
 -1- FUNCTION: May play a role in plant defense. Has probably no  
 oxalate oxidase activity even if the active site is conserved.  
 CC  
 -1- SUBUNIT: Oligomer (believed to be a pentamer but probably hexamer)  
 CC  
 (By similarity).  
 CC  
 -1- SUBCELLULAR LOCATION: Apoplast (By similarity).  
 CC  
 -1- SIMILARITY: Belongs to the germin family.  
 CC  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 -----  
 CC EMBL; AC016829; AAF26795.1; -;  
 DR HSP; P45850; 1F12;  
 DR InterPro; IPR006045; Cupin.  
 DR InterPro; IPR007113; Cupin\_sup.  
 DR InterPro; IPR001929; Germin.  
 DR Pfam; PF00190; Cupin; 1.  
 DR PRINTS; PR00325; GERMIN.  
 DR PROSITE; PS00725; GERMIN; FALSE NEG.  
 KW Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;  
 Multigene family.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 222 GERMIN-LIKE PROTEIN SUBFAMILY 1 MEMBER 4.  
 FT METAL 112 112 MANGANESE (BY SIMILARITY).  
 FT METAL 114 114 MANGANESE (BY SIMILARITY).  
 FT SITE 119 119 PROBABLE NON-FUNCTIONAL MANGANESE-BINDING  
 SITE.  
 FT METAL 161 161 MANGANESE (BY SIMILARITY).  
 FT DISULFID 34 50 BY SIMILARITY.

FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 222 AA; 24102 MW; 036F2D1F6B6C3537 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 155 LAKIIL 160  
 DB 9 LAKIIL 14  
 RESULT 135  
 GL15 ARATH STANDARD; PRT; 222 AA.  
 ID GL15 ARATH  
 AC Q9M8X5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Germin-like protein subfamily 1 member 5 precursor.  
 GN AT3G04190 OR T6K12.19.  
 OS Arabidopsis thaliana (Mouse-ear cross).  
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC euroids II; Brassicales; Brassicaceae; Arabidopsids.  
 CC NCBI\_TaxID=3702;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=cv. Columbia;  
 CC MEDLINE=21016720; PubMed=11130713;  
 CC Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unefeld M.,  
 Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,  
 Delany M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,  
 De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,  
 Wincker P., Cattolico L., Weissbach J., Saurin W., Quétier F.,  
 Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 Wurmback E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 Vezi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,  
 Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,  
 Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Clement J.,  
 Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,  
 de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
 Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
 Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
 Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,  
 Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,  
 Sasaoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
 Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana";  
 Nature 408:820-822(2000).  
 -1- FUNCTION: May play a role in plant defense. Has probably no  
 oxalate oxidase activity even if the active site is conserved.  
 CC  
 -1- SUBUNIT: Oligomer (believed to be a pentamer but probably hexamer)  
 CC  
 (By similarity).  
 CC  
 -1- SUBCELLULAR LOCATION: Apoplast (By similarity).  
 CC  
 -1- SIMILARITY: Belongs to the germin family.  
 CC  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 -----  
 CC EMBL; AC016829; AAF26795.1; -;  
 DR HSP; P45850; 1F12;  
 DR InterPro; IPR006045; Cupin.  
 DR InterPro; IPR007113; Cupin\_sup.  
 DR InterPro; IPR001929; Germin.  
 DR Pfam; PF00190; Cupin; 1.  
 DR PRINTS; PR00325; GERMIN.  
 DR PROSITE; PS00725; GERMIN; FALSE NEG.  
 KW Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;  
 Multigene family.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 222 GERMIN-LIKE PROTEIN SUBFAMILY 1 MEMBER 4.  
 FT METAL 112 112 MANGANESE (BY SIMILARITY).  
 FT METAL 114 114 MANGANESE (BY SIMILARITY).  
 FT SITE 119 119 PROBABLE NON-FUNCTIONAL MANGANESE-BINDING  
 SITE.  
 FT METAL 161 161 MANGANESE (BY SIMILARITY).  
 FT DISULFID 34 50 BY SIMILARITY.

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

-----

EMBL; AC016829; AAR26794.1; -  
HSSP; P45850; 1F12.  
InterPro; IPR006045; Cupin.  
InterPro; IPR007113; Cupin sup.  
InterPro; IPR001929; Germin.  
Pfam; PF00190; Cupin; 1.  
PRINTS; PR00325; GERMIN.  
PROSITE; PS00725; GERMIN; 1.  
Apolplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;  
Multigene family.

FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 222 GERMIN-LIKE PROTEIN SUBFAMILY 1 MEMBER 5.  
FT METAL 112 112 MANGANESE (BY SIMILARITY).  
FT METAL 114 114 MANGANESE (BY SIMILARITY).  
FT METAL 119 119 MANGANESE (BY SIMILARITY).  
FT METAL 163 163 MANGANESE (BY SIMILARITY).  
FT DISULFID 34 50 BY SIMILARITY.  
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 222 AA; 24051 MW; 334DBBF1BAA4F0A8 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160  
Db 9 LAKIIL 14

-----

RESULT 136  
RPC1 BPD3 STANDARD; PRT; 223 AA.  
AC Q37906;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Repressor protein CI.  
GN CI.  
OS Bacteriophage D3.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
OC Lambda-like viruses.  
OK NCBI\_TaxID=31535;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94327444; PubMed=8050993;  
RA Farinha M.A., Allan B.J., Gertman E.M., Ronald S.L., Kropinski A.M.;  
RT "Cloning of the early promoters of Pseudomonas aeruginosa  
bacteriophage D3: sequence of the immunity region of D3."  
J. Bacteriol. 176:4809-4815 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20485557; PubMed=11029426;  
RA Kropinski A.M.;  
RT "Sequence of the genome of the temperate, serotype-converting,  
Pseudomonas aeruginosa bacteriophage D3."  
J. Bacteriol. 182:6066-6074 (2000).  
RN [1]  
RP SIMILARITY: Contains 1 HTH cro/CI-type DNA-binding domain.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF165214; AAA53553.1; -  
HSSP; P16117; 1R69.  
InterPro; IPR001387; HTH\_3.

Pfam; PF01381; HTH\_3; 1.  
SMART; SM00530; HTH\_XRE; 1.  
PROSITE; PS0943; HTH\_CROCI; 1.  
KW Repressor; Transcription regulation; DNA-binding.  
FT DOMAIN 7 60 HTH CRO/CI-TYPE.  
FT DNA BIND 18 37 H-T-H MOTIF (POTENTIAL).  
SQ SEQUENCE 223 AA; 24558 MW; B74F89FE0EB3E781 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 223;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SKPENL 255  
Db 122 SKPENL 127

-----

RESULT 137  
PLF1 MOUSE STANDARD; PRT; 224 AA.  
ID PLF1\_MOUSE  
AC P04095;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 01-NOV-1986 (Rel. 03, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Proliferin 1 precursor (Mitogen-regulated protein 1).  
GN PLF OR PLF1 OR MRP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OK NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=84272617; PubMed=6087314;  
RA Linzer D.I.H., Nathans D.;  
RT "Nucleotide sequence of a growth-related mRNA encoding a member of  
the prolactin-growth hormone family."  
Proc. Natl. Acad. Sci. U.S.A. 81:4255-4259 (1984).  
RN [2]  
RP SEQUENCE OF 1-10 FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=88029317; PubMed=3478191;  
RA Linzer D.I.H., Mordacq J.C.;  
RT "Transcriptional regulation of proliferin gene expression in response  
to serum in transfected mouse cells."  
EMBO J. 6:2281-2288 (1987).  
CC  
CC -!- FUNCTION: May have a role in embryonic development. It is  
CC likely to provide a growth stimulus to target cells in maternal  
CC and fetal tissues during the development of the embryo at mid-  
CC gestation.  
CC  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC  
CC -!- SIMILARITY: Belongs to the somatotropin/prolactin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X05787; CAA29231.1; -  
EMBL; X02245; AAA39946.1; -  
EMBL; X05786; CAA29230.1; -  
PIR; A05086; A05086.  
HSP; Q28632; IAN3.  
MGD; MGI:97618; Plf.  
InterPro; IPR001400; Somatotropin.  
Pfam; PF00103; hormone; 1.  
PRINTS; PR00836; SOMATOTROPIN.  
PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
Hormone; Signal; Multigene family.

```

FT SIGNAL 1 29
FT CHAIN 30 224
FT DISULFID 33 40
FT DISULFID 87 199
FT DISULFID 216 224
FT SEQUENCE 224 AA; 25367 MW; 3786F100C338374B CRC64;

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 224;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LRDKSP 267
|||||
Db 79 LRDKSP 84

RESULT 138
PLF2 MOUSE
ID PLF2 MOUSE STANDARD; PRT; 224 AA.
AC P04768;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Proliferin 2 precursor (Mitogen-regulated protein 2).
GN PLF2 OR MRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=85242683; PubMed=3859868;
RA Linzer D.I.H., Lee S.-J., Ogren L., Talamantes F., Nathans D.;
RT Identification of proliferin mRNA and protein in mouse placenta.;
RL Proc. Natl. Acad. Sci. U.S.A. 82:4356-4359(1985).
CC -1- FUNCTION: May have a role in embryonic development. It is
CC likely to provide a growth stimulus to target cells in maternal
CC and fetal tissues during the development of the embryo at mid-
CC gestation.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the somatotropin/prolactin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; K03235; AAA39945.1; -.
CC HSP; Q28632; IAN3.
CC MGD; MGI:1341833; Plf2.
CC InterPro; IPR001400; Somatotropin.
CC PRINTS; PR00836; Somatotropin.
CC PROSITE; PS00266; SOMATOTROPIN 1; 1.
CC PROSITE; PS00338; SOMATOTROPIN 2; 1.
CC Hormone; Signal; Multigene family.
CC SIGNAL 1 29
CC CHAIN 30 224
CC DISULFID 33 40
CC DISULFID 87 199
CC DISULFID 216 224
CC SEQUENCE 224 AA; 25312 MW; 1EB34BEA21433B82 CRC64;

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 224;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LRDKSP 267
|||||

```

```

Db 79 LRDKSP 84

RESULT 139
PLF3 MOUSE
ID PLF3 MOUSE STANDARD; PRT; 224 AA.
AC P18918;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE Proliferin 3 precursor (Mitogen-regulated protein 3).
GN MRPLF3 OR PLF3 OR MRP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Fibroblast;
RX MEDLINE=90001249; PubMed=2790033;
RA Connor A.M., Waterhouse P., Khokha R., Denhardt D.T.;
RT "Characterization of a mouse mitogen-regulated protein/proliferin
RT gene and its promoter: a member of the growth hormone/prolactin gene
RT superfamily.";
RL Biochim. Biophys. Acta 1009:75-82(1989).
RN [2]
RP SEQUENCE OF 208-224 FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=94319082; PubMed=8043949;
RA Ko M.-S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,
RA Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL Mamm. Genome 5:349-355(1994).
CC -1- FUNCTION: May have a role in embryonic development. It is
CC likely to provide a growth stimulus to target cells in maternal
CC and fetal tissues during the development of the embryo at mid-
CC gestation.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the somatotropin/prolactin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X16009; CAA34146.1; -.
CC EMBL; X16010; CAA34146.1; JOINED.
CC EMBL; X16011; CAA34146.1; JOINED.
CC EMBL; X16012; CAA34146.1; JOINED.
CC EMBL; X16013; CAA34146.1; JOINED.
CC EMBL; U05747; AAB60482.1; -.
CC PIR; S05648; S05648.
CC HSP; Q28632; IAN3.
CC MGD; MGI:1347041; Mrplf3.
CC InterPro; IPR001400; Somatotropin.
CC Pfam; PF00103; hormone; 1.
CC PRINTS; PR00836; SOMATOTROPIN.
CC PROSITE; PS00266; SOMATOTROPIN 1; 1.
CC PROSITE; PS00338; SOMATOTROPIN 2; 1.
CC Hormone; Signal; Multigene family.
CC SIGNAL 1 29
CC CHAIN 30 224
CC DISULFID 33 40
CC DISULFID 87 199
CC DISULFID 216 224
CC SEQUENCE 224 AA; 25338 MW; C87F3A2310C91320 CRC64;

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 224;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 262 LRDKSP 267  
Db 79 LRDKSP 84

RESULT 140  
AROD HALN1 STANDARD; PRT; 226 AA.  
AC QHSB4;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 3-dehydroquininate dehydratase (EC 4.2.1.10) (3-dehydroquinase) (Type I  
DE DHQase)  
GN AROD OR VNG0314G.  
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Shrogha J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlshocher M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
CC -1- CATALYTIC ACTIVITY: 3-dehydroquininate = 3-dehydroshikimate + H(2)O.  
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
CC third step.  
CC -1- SIMILARITY: Belongs to the type-I 3-dehydroquinase family.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; AE004992; AAG18893.1; ALT\_INIT.  
DR HAMAP; MF 00214; -; 1.  
DR InterPro; IPR001381; DHQuinase I.  
DR Pfam; PF01487; DHQuinase I; 1.  
DR ProDom; PD005337; DHQuinase I; 1.  
DR PROSITE; PS01028; DHYDROQUINASE\_I; FALSE NEG.  
KW Aromatic amino acid biosynthesis; Lyase; Complete proteome.  
FT ACT SITE 120 120 BY SIMILARITY.  
FT ACT SITE 146 146 SCHIFF-BASE INTERMEDIATE (BY SIMILARITY).  
SQ SEQUENCE 226 AA; 23048 MW; 0054520F1EF93468 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 226;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 LIADLQ 87  
Db 221 LIADLQ 226

RESULT 141  
RADC STAEF STANDARD; PRT; 226 AA.  
AC Q8CNZ4;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE DNA repair protein radC homolog.  
GN RADC OR SE1336.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RX PubMed=12950922;  
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
RA Qian Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;  
RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
RT Staphylococcus epidermidis strain (ATCC 12228).";  
RL Mol. Microbiol. 49:1577-1593(2003).  
CC -1- FUNCTION: Involved in DNA repair (By similarity).  
CC -1- SIMILARITY: Belongs to the radC family.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; AE016748; AA004935.1; -;  
DR HAMAP; MF 00018; -; 1.  
DR InterPro; IPR001405; RadC.  
DR Pfam; PF04002; RadC; 1.  
DR ProDom; PD007415; RadC; 1.  
DR TIGRFAMs; TIGR00608; RadC; 1.  
DR PROSITE; PS01302; RADC; 1.  
KW DNA repair; Complete proteome.  
SQ SEQUENCE 226 AA; 25319 MW; 16DE9D290C98F005 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 226;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 IRDLKK 334  
Db 57 IRDLKK 62

RESULT 142  
VATE MANSE STANDARD; PRT; 236 AA.  
AC P31402;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Vacuolar ATP synthase subunit E (EC 3.6.3.14) (V-ATPase E subunit)  
DE (Vacuolar proton pump E subunit) (V-ATPase 28 kDa subunit).  
GN VHA26.  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Sphingioidea;  
OC Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Midgut;  
RX MEDLINE=94153976; PubMed=8110816;  
RA Graef R., Harvey W.R., Wiczorek H.;  
RT "Cloning, sequencing and expression of cDNA encoding an insect  
RT V-ATPase subunit E.";  
RL Biochim. Biophys. Acta 1190:193-196(1994).  
CC -1- FUNCTION: The vacuolar ATPase is responsible for acidifying a  
CC variety of intracellular compartments in eukaryotic cells. This  
CC subunit is essential for its assembly or catalytic function. V-  
CC ATPase is responsible for energizing electrophoretic K+/2H+



CC antiport by generating a transmembrane voltage of more than 200mV.  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
 CC H(+) (out).  
 CC -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a  
 CC peripheral catalytic V1 complex (components A to H) attached to  
 CC an integral membrane V0 proton pore complex (components: a, c, c',  
 CC c'', and d).  
 CC -!- SIMILARITY: Belongs to the V-ATPase E subunit family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X67131; CRAA47610.1; -;  
 CC InterPro; IPR002842; ATPsynth\_Esub.  
 CC Pfam; PF01991; V-ATP-synt\_E; 1.  
 CC Hydrolyase; ATP synthetase; Hydrogen ion transport.  
 CC SEQUENCE 226 AA; 26090 MW; DFTAB42E01067400 CRC64;  
 CC  
 CC Query Match 1.8%; Score 6; DB 1; Length 226;  
 CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 78 LLVTLI 83  
 CC |||||  
 CC DB 115 LLVTLI 120  
 CC  
 CC RESULT 143  
 CC GL13 ARATH STANDARD; PRT; 227 AA.  
 CC AC Q9M8X3;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Germin-like protein subfamily 1 member 3 precursor.  
 CC GN AT3G04170 OR T6K12.21.  
 CC OS Arabidopsis thaliana (Mouse-ear cress).  
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CC OX NCBI\_TaxID=3702;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=cv. Columbia;  
 CC RX MEDLINE=21016720; PubMed=11130713;  
 CC RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M., Obermaier B.,  
 CC Farmann B., Vallie G., Bloeker H., Perez-Alonso M., Obermaier B.,  
 CC Deiseny M., Boutry M., Griwell L.A., Mache R., Puigdomenech P.,  
 CC De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,  
 CC Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,  
 CC Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 CC Wurmbach E., Drzonek H., Exfle H., Jordan N., Bangert S.,  
 CC Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 CC Verzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,  
 CC Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,  
 CC Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,  
 CC Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 CC Cooke R., Laudie M., Berger-Llauro C., Fumelle B., Masuy D.,  
 CC de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 CC Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
 CC Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 CC Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 CC Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,  
 CC Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 CC Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,  
 CC Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
 CC Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 CC Sasamoto S., Kimura T., Idegawa K., Kawashima K., Kishida Y.,

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:820-822(2000).  
 CC -!- FUNCTION: May play a role in plant defense. Has probably no  
 CC oxalate oxidase activity even if the active site is conserved.  
 CC -!- SUBUNIT: Oligomer (believed to be a pentamer but probably hexamer)  
 CC (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Apoplast (By similarity).  
 CC -!- SIMILARITY: Belongs to the germin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AC016829; AAF26796.1; -;  
 CC HSP; P45850; IP12.  
 CC InterPro; IPR006045; Cupin.  
 CC InterPro; IPR007113; Cupin\_sup.  
 CC InterPro; IPR001929; Germin.  
 CC Pfam; PF00130; Cupin; 1.  
 CC PRINTS; PR00325; GERMIN.  
 CC PROSITE; PS00725; GERMIN; 1.  
 CC Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;  
 CC Multigene family.  
 CC SIGNAL 1 24 POTENTIAL.  
 CC CHAIN 25 227 GERMIN-LIKE PROTEIN SUBFAMILY 1 MEMBER 3.  
 CC METAL 109 109 MANGANESE (BY SIMILARITY).  
 CC METAL 111 111 MANGANESE (BY SIMILARITY).  
 CC METAL 116 116 MANGANESE (BY SIMILARITY).  
 CC METAL 160 160 MANGANESE (BY SIMILARITY).  
 CC DISULFID 34 50 BY SIMILARITY.  
 CC CARBOHYD 136 136 N-LINKED (GLCNAC ...) (POTENTIAL).  
 CC SEQUENCE 227 AA; 24424 MW; C42948B9F92F9304 CRC64;  
 CC  
 CC Query Match 1.8%; Score 6; DB 1; Length 227;  
 CC Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 155 LAKIIL 160  
 CC |||||  
 CC DB 9 LAKIIL 14  
 CC  
 CC RESULT 144  
 CC Y685\_METH STANDARD; PRT; 232 AA.  
 CC ID Y685\_METH  
 CC AC 026781;  
 CC DT 30-MAY-2000 (Rel. 39, Created)  
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Hypothetical protein MTH685.  
 CC GN MTH685.  
 CC OS Methanobacterium thermoautotrophicum.  
 CC OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 CC OC Methanobacteriaceae; Methanothermobacter.  
 CC OX NCBI\_TaxID=187420;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=Delta H;  
 CC RX MEDLINE=98037514; PubMed=9371463;  
 CC RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 CC Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 CC Harrison D., Hoang L., Keagle P., Lum W., Pochier B., Qiu D.,  
 CC Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 CC Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 CC McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,



RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- SIMILARITY: Belongs to the UPF0023 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE000848; AAB85130.1; - -  
 DR F1R; C69191; C69191.  
 DR InterPro; IPR003022; EFG\_III.V.  
 DR InterPro; IPR002140; UPF0023.  
 DR Pfam; PF01172; UPF0023; 1.  
 DR ProDom; PD003796; UPF0023; 1.  
 DR TIGRFAMs; TIGR00291; TIGR00291; 1.  
 DR PROSITE; PS01267; UPF0023; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 232 AA; 26193 MW; 80ACC4306A5BF1AF CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 36 DKASEE 41  
 Db 57 DKASEE 62  
 RESULT 145  
 ID PSB1 DROME STANDARD; PRT; 235 AA.  
 AC P40304; Q9VVB4;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Proteasome subunit beta type 1 (EC 3.4.25.1) (Proteasome 26 kDa  
 DE subunit).  
 GN PROS26 OR PROS-26 OR L(3)73AI OR CG4097.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94022270; PubMed=8415617;  
 RA Saville K.J., Belote J.M.;  
 RT "Identification of an essential gene, l(3)73AI, with a dominant  
 RT temperature-sensitive lethal allele, encoding a Drosophila proteasome  
 RT subunit".  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8842-8846(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=Berkeley;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*".  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: The proteasome is a multicatalytic proteinase complex  
 CC which is characterized by its ability to cleave peptides with Arg,  
 CC phe, Tyr, Ieu, and Glu adjacent to the leaving group at neutral or  
 CC slightly basic pH. The proteasome has an ATP-dependent proteolytic  
 CC activity.  
 CC -1- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad  
 CC specificity.  
 CC -1- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal  
 CC proteolytic pathway.  
 CC -1- SUBUNIT: The proteasome is composed of at least 15 non identical  
 CC subunits which form a highly ordered ring-shaped structure (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U00790; AAC46465.1; - -  
 DR EMBL; AE003526; AAP49435.1; - -  
 DR EMBL; AY051697; AAK93121.1; - -  
 DR HSPSP; P23724; IRYP.  
 DR MEROPS; T01.986; - -  
 DR FlyBase; FBgn002284; Pros26.  
 DR InterPro; IPR000243; Pept\_T1A\_subB.  
 DR InterPro; IPR001353; Peptidase\_T1.  
 DR Pfam; PF00227; proteasome; 1.  
 DR PROSITE; PS00854; PROTEASOME\_B; 1.  
 KW Proteasome; Hydrolase; Protease; Threonine protease.  
 FT CONFLICT 192 192 R -> W (IN REF. 1).  
 SQ SEQUENCE 235 AA; 25842 MW; 592307AAEDD98B25 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 235;



```

RN  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX  MEDLINE=21359325; PubMed=1146286;
RA  Neolling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA  Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA  Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA  Bennett G.N., Koonin E.V., Smith D.R.;
RT  "Genome sequence and comparative analysis of the solvent-producing
RT  bacterium Clostridium acetobutylicum.";
RL  J. Bacteriol. 183:4923-4938 (2001).
CC  -!- CATALYTIC ACTIVITY: ATP + 7,8-diaminononanoate + CO(2) = ADP +
CC  phosphate + dethiobiotin.
CC  -!- COFACTOR: Magnesium (By similarity).
CC  -!- PATHWAY: Bioconversion of pimelate into dethiobiotin.
CC  -!- SIMILARITY: Belongs to the dethiobiotin synthetase family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  PIR; AE007647; AAK79329.1; -.
DR  HANAP; MF 00336; -.
DR  InterPro; IPR004472; BioD_synth.
DR  InterPro; IPR002586; CbiA_P.
DR  Pfam; PF01656; CbiA; 1.
DR  TIGRFAMs; TIGR00347; bioD; 1.
KW  Biotin biosynthesis; Ligase; Magnesium; ATP-binding;
KW  Complete proteome.
FT  NP_BIND 8 ATP (BY SIMILARITY).
SQ  SEQUENCE 240 AA; 26677 MW; 2FA903A8C23E4AD1 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 EDYEKL 216
Db 66 EDYEKL 71

RESULT 149
Y538_VIBPA STANDARD; PRT; 242 AA.
AC Q87S86;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Hypothetical response regulatory protein VP0538.
GN VP0538.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=13620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
CC -!- SIMILARITY: Contains 1 HTH LytR-type DNA-binding domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AP005074; BAC58801.1; -.
DR  InterPro; IPR007492; LytTR.
DR  InterPro; IPR001789; Response_reg.
DR  Pfam; PF04397; LytTR; 1.
DR  Pfam; PF00072; response_reg; 1.
DR  ProDom; PD000039; Response_reg; 1.
DR  SMART; SM00448; REC; 1.
DR  PROSITE; PS09330; HTH_LYTTR; 1.
DR  PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT DOMAIN 1 116 RESPONSE_REGULATORY.
FT DOMAIN 139 240 HTH_LYTTR-TYPE.
SQ SEQUENCE 242 AA; 26992 MW; A7264E539FID4CD7 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 DDEQFA 319
Db 8 DDEQFA 13

RESULT 150
GLPF_MYCGA STANDARD; PRT; 243 AA.
AC P52280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable glycerol uptake facilitator protein.
GN GLPF OR MYCGA0140 OR MGA_0641.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Forsyth M.H., Saoud S., Geary S.J.;
RT "Gene encoding a glycerol uptake protein from Mycoplasma gallisepticum,
RT strain S6.";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RX STRAIN=R(1ow);
RX MEDLINE=22830409; PubMed=12949158;
RA Papazisi L., Gorton T.S., Kutish G., Markham P.F., Browning G.F.,
RA Nguyen D.K., Swartzell S., Madan A., Mahairas G., Geary S.J.;
RT "The complete genome sequence of the avian pathogen Mycoplasma
RT gallisepticum strain R(1ow).";
RL Microbiology 149:2307-2316 (2003).
CC -!- FUNCTION: Glycerol enters the cell via the glycerol diffusion
CC facilitator protein. This membrane protein facilitates the
CC movement of glycerol across the cytoplasmic membrane (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- DOMAIN: Aquaporins contain two tandem repeats each containing
CC three membrane-spanning domains and a pore-forming loop with the
CC signature motif Asn-Pro-Ala (NPA).
CC -!- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; U35010; AAA79047.1; -.  
 DR EMBL; AE016967; AAP56364.1; ALT\_INIT.  
 DR HSSP; PL1244; 1FX8.  
 DR InterPro; IPR000425; MTP.  
 DR PRINTS; PR00783; MINTRINSCP.  
 DR PRODOM; PD000295; MTP\_family; 1.  
 DR PROSITE; PS00221; MTP; 1.  
 KW Glycylol metabolism; Transport; Repeat; Transmembrane;  
 KW Complete proteome.  
 FT TRANSMEM 7 27 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 88 108 POTENTIAL.  
 FT TRANSMEM 143 163 POTENTIAL.  
 FT TRANSMEM 166 186 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT SITE 72 74 NPA 1.  
 FT SITE 187 189 NPA 2.  
 FT CONFLICT 28 32 S -> C (IN REF. 1).  
 FT CONFLICT 32 32 K -> E (IN REF. 1).  
 FT CONFLICT 48 48 L -> F (IN REF. 1).  
 FT CONFLICT 70 70 N -> H (IN REF. 1).  
 FT CONFLICT 204 243 AFNKLQNPVSADFRYGLVPLLAIAAGLIMGGSLLINQ  
 FT -> VV (IN REF. 1).  
 SQ SEQUENCE 243 AA; 26194 MW; 43594672BCE4990C CRC64;

Query Match 1.8%; Score 6; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LLGELI 236  
 Db 8 LLGELI 13

RESULT 151  
 YG87\_MYCPN STANDARD; PRT; 250 AA.  
 AC Q50315;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein MPN687 (K05\_orf250).  
 GN MPN687 OR MP155.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=96177562; PubMed=8604303;  
 RA Hilbert H., Himmelsreich R., Plagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a  
 RT cluster of ribosomal protein genes";  
 RL Nucleic Acids Res. 24:628-639 (1996).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae";  
 RL Nucleic Acids Res. 24:4420-4449 (1996).  
 [3]  
 RP IDENTIFICATION BY MASS SPECTROMETRY.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=21089919; PubMed=11271496;  
 RA Regula J.T., Ueberle B., Boguth G., Goerg A., Schnoelzer M.,

RA Herrmann R., Frank R.;  
 RT "Towards a two-dimensional proteome map of Mycoplasma pneumoniae";  
 RL Electrophoresis 21:3765-3780 (2000).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U34816; AAC43647.1; -.  
 DR EMBL; AE000017; AAB95803.1; -.  
 DR PIR; S62838; S62838.  
 KW Complete proteome.  
 SQ SEQUENCE 250 AA; 28662 MW; 4378A2834BBB7877 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 287 TQPIVE 292  
 Db 153 TQPIVE 158

RESULT 152  
 ATP6\_MYCLE STANDARD; PRT; 251 AA.  
 AC P45829;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
 GN ATPB OR ML1139.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smith D.R., Robison K.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus";  
 RL Nature 409:1007-1011 (2001).

CC -I- FUNCTION: Key component of the proton channel; it may play a  
 CC direct role in the translocation of protons across the membrane.  
 CC -I- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -I- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 CC core - and CF(0) - the membrane proton channel. CF(1) has five  
 CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)  
 CC has three main subunits: a, b and c.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -I- SIMILARITY: Belongs to the ATPase A chain family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; U15186; AAA63110.1; --  
CC EMBL; AL583920; CAC31520.1; --  
CC PIR; T09980; T09980.  
CC HSP; P00855; ICL7.  
CC Leproma; MLI139; --  
CC InterPro; IPR000568; ATPsynt\_Asub.  
CC Pfam; PF00119; ATP-synt\_A; 1.  
CC PRINTS; PRO0123; ATPASEA.  
CC TIGRFAMS; TIGR01131; ATP synt 6 or A; 1.  
CC PROSITE; PS00449; ATPASE\_A; 1.  
CC Hydrogen ion transport; CF(0); Transmembrane; Complete proteome.  
CC TRANSMEM 48  
CC TRANSMEM 28  
CC TRANSMEM 84 104 POTENTIAL.  
CC TRANSMEM 130 150 POTENTIAL.  
CC TRANSMEM 154 174 POTENTIAL.  
CC TRANSMEM 192 212 POTENTIAL.  
CC TRANSMEM 220 240 POTENTIAL.  
CC SEQUENCE 251 AA; 27558 MW; BCF2AE7FC158DF9 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLIA 84

DB 198 LVTLIA 203

RESULT 153

ID TPIS\_COXBU STANDARD; PRT; 255 AA.  
AC Q83BQ3;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).  
GN TPIA OR CBUL450.  
OS Coccidia burnetii.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;  
OC Coccidiaceae; Coccidia.  
OX NCBI\_TaxID=777;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nine Mile phase I / RSA 493;  
RX MEDLINE=22608657; PubMed=12704232;  
RA Seshadri R., Paulsen I.T., Edeen J.A., Read T.D., Nelson K.E.,  
RA Nelson W.C., Ward N.L., Tettelin H., Davidson T.M., Beanan M.J.,  
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,  
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,  
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;  
RT "Complete genome sequence of the Q-fever pathogen, Coccidia  
burnetii.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003)  
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone  
phosphate.  
CC -1- PATHWAY: Plays an important role in several metabolic pathways.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SIMILARITY: Belongs to the triosephosphate isomerase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE016964; AA090947.1; --

DR TIGR; CBUL450; --  
DR HAMAP; MF\_00147; --; 1.  
DR InterPro; IPR000652; Triophos\_ismrse.  
DR Pfam; PF00121; TIM; 1.  
DR ProDom; PD001005; Triophos\_ismrse; 1.  
DR TIGRFAMS; TIGR00419; tim; 1.  
DR PROSITE; PS00171; TIM; 1.  
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;  
KW Pentose shunt; Complete proteome.  
FT ACT\_SITE 96  
FT ACT\_SITE 169 169 BY SIMILARITY.  
FT ACT\_SITE 169 169 BY SIMILARITY.  
SQ SEQUENCE 255 AA; 28172 MW; 82780F6A8985D2C6 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 255;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 ALRCGI 144

DB 117 ALRCGI 122

RESULT 154

ID COAT\_ICMV STANDARD; PRT; 256 AA.  
AC Q08583;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Coat protein.  
DE Coat protein.  
GN ARI.  
OS Indian cassava mosaic virus (ICMV).  
OC Viruses; dsDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=31600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94065670; PubMed=8245859;  
RA Hong Y.G., Robinson D.J., Harrison B.D.;  
RT "Nucleotide sequence evidence for the occurrence of three distinct  
RT whitefly-transmitted geminiviruses in cassava.";  
RA J. Gen. Virol. 74:2437-2443(1993).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z24758; CAA80885.1; --

DR PIR; JQ2326; JQ2326.

DR InterPro; IPR000650; Gem\_coat\_ARI.

DR InterPro; IPR000263; GV\_A/Bri\_coat.

DR Pfam; PF00844; Gemini\_coat; 1.

DR PRINTS; PR00223; GEMCOATARBRI.

DR ProDom; PD000901; Gem\_coat\_ARI; 1.

KW Coat protein.

SQ SEQUENCE 256 AA; 29910 MW; 0FA4FD4AE48A702A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 256;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 VTKRQS 228

DB 38 VTKRQS 43

RESULT 155

PCNA\_NPVAC

ID PCNA\_NPVAC STANDARD; PRT; 256 AA.

AC P11038;

01-JUL-1989 (Rel. 11, Created)  
 01-JUL-1989 (Rel. 11, Last sequence update)  
 15-JUL-1998 (Rel. 36, Last annotation update)  
 Proliferating cell nuclear antigen (EcoRI-T site protein ETL).  
 PCNA OR ETL.  
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID=46015;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=L1;  
 RX MEDLINE=88275045; PubMed=3292791;  
 RA Crawford A., Miller L.K.;  
 RT "Characterization of an early gene accelerating expression of late  
 genes of the baculovirus Autographa californica nuclear polyhedrosis  
 virus.";  
 RL J. Virol. 62:2773-2781(1988).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C6;  
 RX MEDLINE=94303173; PubMed=8030224;  
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;  
 RT "The complete DNA sequence of Autographa californica nuclear  
 polyhedrosis virus.";  
 RL Virology 202:586-605(1994).  
 [3]  
 RN SEQUENCE OF 1-126 FROM N.A.  
 RC STRAIN=L1;  
 RX MEDLINE=94267933; PubMed=8207843;  
 RA Passarelli A.L., Todd J.W., Miller L.K.;  
 RT "A baculovirus gene involved in late gene expression predicts a large  
 polypeptide with a conserved motif of RNA polymerases.";  
 RL J. Virol. 68:4673-4678(1994).  
 [4]  
 RN SIMILARITY TO PCNA, AND FUNCTION.  
 RX MEDLINE=69143717; PubMed=2563897;  
 RA O'Reilly D.R., Crawford A.M., Miller L.K.;  
 RT "Viral Proliferating cell nuclear antigen.";  
 RL Nature 337:606-606(1989).  
 CC -!- FUNCTION: REQUIRED FOR TIMELY EXPRESSION OF MANY OTHER ACNPV  
 GENES.  
 CC -!- DEVELOPMENTAL STAGE: WAS FOUND TO BE TRANSCRIBED EARLY (FROM 2  
 THROUGH 6 HOURS POST-INFECTION) BUT NOT LATE (12 TO 24 HOURS POST-  
 INFECTION) IN INFECTION.  
 CC -!- SIMILARITY: Belongs to the PCNA family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M20718; AAA21097.1; -;  
 CC EMBL; L22858; AAA66679.2; -;  
 CC EMBL; U04879; AAA20059.1; -;  
 CC PIR; A28147; WMNVET.  
 CC HSSP; P12004; IAXC.  
 CC InterPro; IPR000730; Pr\_cel\_nuc\_antig.  
 CC Pfam; PF00705; PCNA; 1.  
 CC Pfam; PF02747; PCNA\_C; 1.  
 CC PRINTS; PR00339; PCNACVCLIN.  
 CC ProDom; PD002673; Pr\_cel\_nuc\_antig; 1.  
 CC TIGRFAMs; TIGR00590; pcna; 1.  
 CC PROSITE; PS00293; PCNA\_2; 1.  
 CC PROSITE; PS01251; PCNA\_1; 1.  
 CC KW DNA-binding; Nuclear protein; DNA replication; Early protein.  
 FT DNA\_BIND 61 80  
 FT SEQUENCE 256 AA; 28635 MW; 73A8A9887AD9E532 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 186 TFXDLL 191  
 DB 18 TFXDLL 23  
 [|||||]  
 RESULT 156  
 Y789 TREPA  
 ID Y789 TREPA STANDARD; PRT; 258 AA.  
 AC Q83768;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein TP0789 precursor.  
 GN TP0789.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Nichols;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
 McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 spirochete.";  
 RL Science 281:375-388(1998).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE001250; AAC65764.1; -;  
 CC PIR; D71280; D71280.  
 CC TIGR; TP0789; -;  
 CC KW Hypothetical protein; Signal; Complete proteome.  
 FT SIGNAL 1 23  
 FT CHAIN 24 258 HYPOTHETICAL PROTEIN TP0789.  
 SQ SEQUENCE 258 AA; 29040 MW; 5CE4E6C09FC240B1 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 301 LIEFLS 306  
 DB 75 LIEFLS 80  
 [|||||]  
 RESULT 157  
 PTB2 ANASP  
 ID PTB2 ANASP STANDARD; PRT; 260 AA.  
 AC Q8YIE2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Phosphate import ATP-binding protein pstB 2 (EC 3.6.3.27) (Phosphate-  
 transporing ATPase 2) (ABC phosphate transporter 2).  
 GN PSTB2 OR ALL0908.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;

RN SEQUENCE FROM N.A.  
 RP MEDLINE=21595285; PubMed=11759840;  
 RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Katanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -1- FUNCTION: Part of the ABC transporter complex pstsABC  
 CC (TC 3.A.1.7.1) involved in phosphate import. Responsible for  
 CC energy coupling to the transport system (By similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + phosphate(Out) = ADP + phosphate  
 CC + phosphate(In).  
 CC -1- SUBUNIT: The complex is composed of two ATP-binding proteins  
 CC (pstB), two transmembrane proteins (pstC and pstA) and a solute-  
 CC binding protein (psts) (Probable).  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).  
 CC -1- SIMILARITY: Belongs to the ABC transporter family. PstB subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; APO03584; BAB72865.1; -;  
 DR PIR; A11919; A11919.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC TRANSPORTER 1; 1.  
 DR PROSITE; PS00893; ABC TRANSPORTER 2; 1.  
 KW Hydrolase; Transport; Phosphate transport; Membrane; Inner membrane;  
 KW ATP-binding; Complete proteome.  
 FT NP BIND 41 48 ATP (By similarity).  
 SQ SEQUENCE 260 AA; 28967 MW; E379845A0320F7F CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 105 LRRQIG 110 PRT; 261 AA.  
 DB 87 LRRQIG 92  
 RESULT 158  
 THIG MYCLE STANDARD; PRT; 261 AA.  
 ID Q9ZBL2;  
 AC Q9ZBL2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Thiazole biosynthesis protein thig.  
 GN THIG OR ML0297 OR MLCB1450.26.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 CX NCBI\_TaxID=1769;  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus";  
 RL Nature 409:1007-1011(2001).  
 CC -1- FUNCTION: Required for the synthesis of the thiazole moiety of  
 CC thiamine (By similarity).  
 CC -1- COFACTOR: FMN (Potential).  
 CC -1- PATHWAY: Thiamine biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the thig family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL035159; CAA22710.1; -;  
 DR EMBL; AL583918; CAC23805.1; -;  
 DR PIR; T44741; T44741.  
 DR Leproma; ML0297; -; 1.  
 DR HAMAP; MF\_00443; -; 1.  
 DR InterPro; IPR003009; FMN enzyme.  
 DR InterPro; IPR008867; Thig.  
 DR DR Pfam; PF05690; Thig; 1.  
 KW Thiamine biosynthesis; Flavoprotein; FMN; Complete proteome.  
 SQ SEQUENCE 261 AA; 27138 MW; A8CC3CAD6F5DE49D CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 24 NLAIIE 29  
 DB 26 NLAIIE 31  
 RESULT 159  
 LEFS NPVOP STANDARD; PRT; 263 AA.  
 ID LEFS\_NPVOP  
 AC O10344;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Late expression factor 5.  
 GN LEF-5.  
 OS Orgyia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 CX NCBI\_TaxID=164623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97271300; PubMed=9126251;  
 RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
 RA Rohrmann G.F.;  
 RT "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear  
 RT polyhedrosis virus genome";  
 RL Virology 229:381-399(1997).  
 CC -1- FUNCTION: Required for late and very late gene expression (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the baculoviruses LEF-5 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U75930; AAC59099.1; -. LEF5.  
 DR InterPro; IPR006923; Baculo\_LEF5.  
 DR Pfam; PF04838; Baculo\_LEF5; 1.  
 KW Transcription regulation.  
 SQ SEQUENCE 263 AA; 30323 MW; 3FE59420A6BD81D4 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIFFL 305  
 |||||  
 DB 42 KLIFFL 47

## RESULT 160

RPOD METH  
 ID RPOD METH STANDARD; PRT; 264 AA.

AC O26144;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE DNA-directed RNA polymerase subunit D (EC 2.7.7.6).  
 GN RPOD OR MTH37.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 ON NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;

RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Adrege T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pochier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiواني N., Caruso A., Bush D., Safer H., Patwell S., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
 CC of DNA into RNA using the four ribonucleoside triphosphates as  
 CC substrates.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC -!- SIMILARITY: Belongs to the archaeobacteria RPOD / eukaryotic RPB3  
 CC RNA polymerase subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE000796; AAB84545.1; -.  
 CC FNR; G69147; G69147.  
 CC HAMAP; MF 00320; -; 1.  
 CC InterPro; IPR009025; RBP11-like RNApo.

CC InterPro; IPR001700; RNA\_pola\_bac.org.  
 CC InterPro; IPR001514; RNA\_pold.  
 CC Pfam; PF01000; RNA\_pola\_bac; 1.  
 CC ProDom; PD002883; RNA\_pold; 1.  
 CC SMART; SM00662; RPOLD; 1.  
 CC PROSITE; PS00446; RNA\_POLD\_30KD; 1.  
 CC Transferrase; DNA-directed RNA polymerase; Transcription;  
 CC Complete proteome.

SQ SEQUENCE 264 AA; 23394 MW; 7740A65424B6AFEB CRC64;

Query Match 1.8%; Score 6; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAOL 69  
 |||||  
 DB 139 EAVAOL 144

## RESULT 161

YJ15 SCHPO  
 ID YJ15 SCHPO STANDARD; PRT; 264 AA.

AC O13679;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C737.05 in chromosome III.  
 GN SPC737.05.  
 GN SPC737.05.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 ON NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellon J., Simmonds S., Squares R., Stevens K., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkhardt G., Aert R., Robben J., Grynolprez B.,  
 RA Welfjens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.I.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Spakovski G.V., Ussery D., Barrall B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AL031546; CAB44773.1; -.  
 CC PIR; T41578; T41578.

CC GeneDB SPombe; SPC737.05; -.  
 CC KW Hypothetical protein; Transmembrane.  
 CC TRANSMEM 48 68 POTENTIAL.  
 CC TRANSMEM 112 132 POTENTIAL.  
 CC TRANSMEM 142 162 POTENTIAL.  
 SQ SEQUENCE 264 AA; 30665 MW; D0E75C409F09F0B CRC64;

Query Match 1.8%; Score 6; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLV 80  
 DB 118 SSGLLV 123  
 |||||

RESULT 162  
 ID 3BH2 MOUSE STANDARD; PRT; 265 AA.  
 AC P26143;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II  
 DE (3beta-HSD II) (Includes: 3-beta-hydroxy-delta(5)-steroid  
 DE dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid  
 DE dehydrogenase) (Progesterone reductase); Steroid delta-isomerase  
 DE (SC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)] (fragment).  
 GN HSD3B2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX BAIN P.A., Yoo M., Clarke T., Hammond S.H., Payne A.H.;  
 RT "Multiple forms of mouse 3 beta-hydroxysteroid dehydrogenase/delta 5-  
 RT delta 4 isomerase and differential expression in gonads, adrenal  
 RT glands, liver, and kidneys of both sexes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8870-8874 (1991).  
 CC -1- FUNCTION: 3beta-HSD is a bifunctional enzyme, that catalyzes the  
 CC oxidative conversion of delta(5)-ene-3-beta-hydroxy steroid, and  
 CC the oxidative conversion of ketosteroids. The 3beta-HSD enzymatic  
 CC system plays a crucial role in the biosynthesis of all classes of  
 CC hormonal steroids.  
 CC -1- CATALYTIC ACTIVITY: 3-beta-hydroxy-delta(5)-steroid + NAD(+) = 3-  
 CC oxo-delta(5)-steroid + NADH.  
 CC -1- CATALYTIC ACTIVITY: A 3-oxo-delta(5)-steroid = a 3-oxo-delta(4)-  
 CC steroid.  
 CC -1- PATHWAY: Steroid biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum and mitochondrial  
 CC membrane-bound protein.  
 CC -1- TISSUE SPECIFICITY: Liver and kidney.  
 CC -1- SIMILARITY: Belongs to the 3beta-HSD family.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL: M75886; -, NOT\_ANNOTATED\_CDS.  
 DR MGD; MGI:96234; Hsd3b2.  
 DR InterPro; IPR002225; 3beta\_HSD.  
 DR Pfam; PF01073; 3beta\_HSD; 1.  
 DR Steroidogenesis; Oxidoreductase; NAD; Isomerase; Mitochondrion;  
 KW Multigene family; Multifunctional enzyme; Transmembrane;  
 KW Endoplasmic reticulum.  
 FT NON TER 1  
 FT TRANSMEM 180 198 POTENTIAL.  
 SQ SEQUENCE 265 AA; 29875 MW; C8074C0C20ABA917 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 KSPNIQ 270  
 DB 137 KSPNIQ 142  
 |||||

RESULT 163  
 ID MCRZ METJA STANDARD; PRT; 266 AA.  
 AC Q60387;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Methyl-coenzyme M reductase II gamma subunit (EC 2.8.4.1) (Coenzyme-B  
 DE sulfotransferase gamma) (MCR II gamma).  
 GN MTRG OR MTRG OR MJO082.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OC NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2651 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8698087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utkerback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073 (1996).  
 CC -1- FUNCTION: Reduction of methyl-coenzyme M (2-(methylthio)  
 CC ethanesulfonic acid) with 7-mercaptoheptanoylthreonine phosphate  
 CC to methane and an heterodisulfide (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 2-(methylthio)ethanesulfonate (methyl-CoM) +  
 CC N-(7-mercaptoheptanoyl)threonine 3-O-phosphate (coenzyme B) = COM-  
 CC S-S-CoB + methane.  
 CC -1- COFACTOR: Binds 2 coenzyme F430 noncovalently per hexamer.  
 CC Coenzyme F430 is a yellow nickel porphyrinoid (By similarity).  
 CC -1- PATHWAY: Methanogenesis; last step.  
 CC -1- SUBUNIT: Hexamer of two alpha, two beta, and two gamma chains (By  
 CC similarity).

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL: U67465; AAB98062.1; -;  
 DR PIR; B64310; B64310.  
 DR HSPG; P11562; LMRO.  
 DR TIGR; MJO082; -;  
 DR InterPro; IPR009024; MCR fer like.  
 DR InterPro; IPR003178; MCR\_gamma.  
 DR Pfam; PF02240; MCR\_gamma; 1.  
 DR PIRSF; PIRSF00264; Meth\_CoM\_gamma; 1.  
 DR ProDom; PD005845; MCR\_gamma; 1.  
 KW Methanogenesis; Transferase; Multigene family; Complete proteome.  
 SQ SEQUENCE 266 AA; 30764 MW; 51E7B3C7F984157A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 EIVKIL 21  
 |||||

```
Db 38 EIVKIL 43

RESULT 164
PSD_HELPY STANDARD; PRT; 267 AA.
ID SUHB_PASMU STANDARD; PRT; 267 AA.
AC Q25911;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphatidylserine decarboxylase proenzyme (EC 4.1.1.65) [Contains:
DE Phosphatidylserine decarboxylase alpha chain; Phosphatidylserine
DE decarboxylase beta chain].
GN PSD OR HP1357.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-P., White O., Kerlavang A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams H.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: Phosphatidyl-L-serine =
CC phosphatidylethanolamine + CO(2).
CC -1- COFACTOR: Pyruvoyl group (By similarity).
CC -1- SIMILARITY: Belongs to the phosphatidylserine decarboxylase
CC family. Subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL; AE000636; AAD08399.1; -.
CC PIR; E64689; E64689.
CC TIGR; HP1357; -.
CC HAMAP; MF_00662; -; 1.
CC InterPro; IPR003817; PS_Dcarboxylase.
CC InterPro; IPR005221; PS_Decarb.
CC Pfam; PF02666; PS_Dcarboxylase; 1.
CC TIGRFAMs; TIGR00163; PS_decarb; 1.
CC Phospholipid biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
CC Complete proteome.
FT CHAIN 1 235 PHOSPHATIDYLSELINE DECARBOXYLASE BETA
FT CHAIN 236 267 CHAIN (BY SIMILARITY).
FT CHAIN 236 267 PHOSPHATIDYLSELINE DECARBOXYLASE ALPHA
FT SITE 235 236 CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).
FT MOD_RES 236 236 CONVERTED TO A PYRUVYL GROUP (BY
FT SIMILARITY).
SQ SEQUENCE 267 AA; 30201 MW; 8CE79AC4126E8A4A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 DKSPNI 269
|||||
```

```
Db 67 DKSPNI 72

RESULT 165
SUHB_PASMU STANDARD; PRT; 267 AA.
ID SUHB_PASMU STANDARD; PRT; 267 AA.
AC Q9CNV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-1-monophosphatase (EC 3.1.3.25) (IMPase) (Inositol-1-
DE phosphatase) [I-1-Pase].
GN SUHB OR PM0315.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida Pm70."
RA Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: Myo-inositol 1-phosphate + H(2)O = myo-
CC inositol + phosphate.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- SIMILARITY: Belongs to the inositol monophosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL; AB006067; AAK02399.1; -.
CC HSSP; P29218; 1IMF.
CC InterPro; IPR000760; Inositol_P.
CC Pfam; PF00459; Inositol_P; 1.
CC ProDom; PD023420; Inositol_P; 1.
CC PROSITE; PS00629; IMP_1; 1.
CC PROSITE; PS00630; IMP_2; 1.
CC Hydrolase; Magnesium; Complete proteome.
SQ SEQUENCE 267 AA; 29470 MW; E7070FAE3BA589CA CRC64;

Query Match 1.8%; Score 6; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 DKASEE 41
|||||
Db 44 DKASEE 49

RESULT 166
143C_ARATH STANDARD; PRT; 268 AA.
ID 143C_ARATH STANDARD; PRT; 268 AA.
AC Q9CSW6; Q9FZD3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 14-3-3-like protein Gp14 iota (General regulatory factor 12).
GN GRF12 OR ATIG26480 OR T1X7.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Alsterfjord M., Rosenquist M., Larsson C., Sommarin M.;
```

RT "Novel 14-3-3 isoforms in Arabidopsis thaliana.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cy. Columbia;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,  
Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,  
Militescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana";  
RL Nature 408:816-820(2000).  
CC -!- FUNCTION: Is associated with a DNA binding complex that binds to  
the G box, a well-characterized cis-acting DNA regulatory element  
found in plants genes (By similarity).  
CC -!- SIMILARITY: Belongs to the 14-3-3 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF335544; AAK11271.1; -;  
DR EMBL; AC013427; AAF98570.1; -;  
DR PIR; F86391; F86391.  
DR HSP; P23312; I338.  
DR InterPro; IPR000308; 14-3-3.  
DR Pfam; PF00244; 14-3-3; 1.  
DR PRINTS; PR00305; 1433ZETA.  
DR PRODOM; PD000600; 14-3-3; 1.  
DR SMART; SM00101; 14-3-3; 1.  
DR PROSITE; PS00796; T433\_1; 1.  
DR PROSITE; PS00797; 1433\_2; 1.  
KW Multigene family.  
KW CONFLICT 268 N -> VVHFPMRTDQRAWKLEI (IN REF. 2).  
SQ SEQUENCE 268 AA; 30545 MW; 914F394CEC07A28C CRC64;  
  
Query Match 1.88; Score 6; DB 1; Length 268;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 130 LKGYEA 135  
Db 155 LKGYEA 160  
|||||  
RESULT 167  
AAKB RAT  
ID AAKB RAT STANDARD; PRT; 269 AA.  
AC P80366; O63048;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 5'-AMP-activated protein kinase, beta-1 subunit (AMPK beta-1 chain)  
DE (AMPKb) (40 kDa subunit).  
GN PKAKB1.  
OS Rattus norvegicus (Rat).  
-----  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RX MEDLINE=96224074; PubMed=8621499;  
RA Gao G., Fernandez C.S., Stapleton D., Auster A.S., Widmer J.,  
Dyck J.R.B., Kemp B.E., Witters L.A.;  
RT "Non-catalytic beta- and gamma-subunit isoforms of the  
5'-AMP-activated protein kinase.";  
RL J. Biol. Chem. 271:8675-8681(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Wistar;  
RX MEDLINE=96215327; PubMed=8626596;  
RA Woods A., Cheung P.C.F., Smith F.C., Davison M.D., Scott J.,  
Beri R.K., Carling D.;  
RT "Characterization of AMP-activated protein kinase beta and gamma  
subunits. Assembly of the heterotrimeric complex in vitro.";  
RL J. Biol. Chem. 271:10282-10290(1996).  
RN [3]  
RP SEQUENCE OF 35-158 FROM N.A., AND SEQUENCE OF 35-71; 78-82 AND 89-158.  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RX MEDLINE=95050763; PubMed=7961907;  
RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.I.,  
Teh T., House C.M., Witters L.A., Kemp B.E.;  
RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are  
homologs of proteins that interact with yeast Snf1 protein kinase.";  
RL J. Biol. Chem. 269:29343-29346(1994).  
RN [4]  
RP PARTIAL SEQUENCE, MYRISTOYLATION, AND PHOSPHORYLATION OF SER-23;  
SER-24; SER-107 AND SER-181.  
RX MEDLINE=97450976; PubMed=9305909;  
RA Mitchell K.I., Michell B.J., House C.M., Stapleton D., Dyck J.,  
Gamble J., Ullrich C., Witters L.A., Kemp B.E.;  
RT "Posttranslational modifications of the 5'-AMP-activated protein  
kinase beta subunit.";  
RL J. Biol. Chem. 272:24475-24479(1997).  
RN [5]  
RP MUTAGENESIS, MYRISTOYLATION, AND PHOSPHORYLATION OF SER-23;  
SER-24; SER-107 AND SER-181.  
RX MEDLINE=21092935; PubMed=11171104;  
RA Warden S.M., Richardson C., O'Donnell J. Jr., Stapleton D., Kemp B.E.,  
Witters L.A.;  
RT "Post-translational modifications of the beta-1 subunit of  
AMP-activated protein kinase affect enzyme activity and cellular  
localization.";  
RL Biochem. J. 354:275-283(2001).  
CC -!- FUNCTION: AMPK IS RESPONSIBLE FOR THE REGULATION OF FATTY ACID  
SYNTHESIS BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. ALSO  
REGULATES CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND  
INACTIVATION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND HORMONE-  
SENSITIVE LIPASE. THIS IS A REGULATORY SUBUNIT, MAY BE A POSITIVE  
REGULATOR OF AMPK ACTIVITY. IT MAY ALSO SERVE AS AN ADAPTER  
MOLECULE FOR THE CATALYTIC ALPHA-SUBUNIT.  
CC -!- SUBUNIT: Heterotrimer of an alpha catalytic subunit, a beta and a  
gamma non-catalytic regulatory subunits  
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney, heart, white  
adipose tissue, lung and spleen.  
CC -!- PTM: Phosphorylated when associated with the catalytic subunit.  
CC -!- SIMILARITY: Belongs to the 5'-AMP-activated protein kinase beta  
subunit family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U42411; AAC52579.1; -;

DR EMBL; X95577; CAA64830.1; --  
 DR InterPro; IPR006828; AMPKBI.  
 DR Pfam; PF04739; AMPKBI.1;  
 KW Fatty acid biosynthesis; Phosphorylation; Myristate; Lipoprotein.  
 FT INIT\_WET 0  
 FT LIPID 1 N-myristoyl glycine.  
 FT MOD RES 23 23 PHOSPHORYLATION (AUTO-).  
 FT MOD RES 24 24 PHOSPHORYLATION (AUTO-).  
 FT MOD RES 107 107 PHOSPHORYLATION (AUTO-).  
 FT MOD RES 181 181 PHOSPHORYLATION.  
 FT CONFLICT 25 25 G -> E (IN REF. 1).  
 FT CONFLICT 51 51 M -> I (IN REF. 3; AA SEQUENCE).  
 SQ SEQUENCE 269 AA; 30263 MW; 6245087E5E81E1 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 3e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 YISKPE 253  
 DB 191 YISKPE 196

## RESULT 168

VG51\_HSVSA STANDARD; PRT; 269 AA.  
 AC Q01036;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Gene 51 glycoprotein.  
 GN 51 OR EDRF2.  
 OS Herpesvirus saimiri (strain 11).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Rhadinovirus.  
 OX NCBI\_TaxID=10383;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92333688; PubMed=1321287;  
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,  
 RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,  
 RA Honess R.W.;  
 RT "Primary structure of the herpesvirus saimiri genome.";  
 RL J. Virol. 66:5047-5058(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9220228; PubMed=1314457;  
 RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;  
 RT "Analysis of nucleotide sequence of the rightmost 43 kbp of  
 RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic  
 RT organization between HVS and Epstein-Barr virus.";  
 RL Virology 188:296-310(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Nicholas J., Coles L.S., Newman C., Honess R.W.;  
 RT "Regulation of the herpesvirus saimiri (HVS) delayed-early  
 RT 110-kilodalton promoter by HVS immediate-early gene products and a  
 RT homolog of the Epstein-Barr virus R trans activator.";  
 RL J. Virol. 62:2457-2466(1988).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X64346; CAA45674.1; --  
 DR EMBL; M86409; AAA46128.1; --  
 DR EMBL; M60850; AAA46160.1; --  
 KW Glycoprotein.  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 53 53

FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 269 AA; 29574 MW; A6038FDE46A8652C CRC64;

Query Match 1.8%; Score 6; DB 1; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 3e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 KTQPIV 291  
 DB 226 KTQPIV 231

## RESULT 169

Y440\_MYCGE STANDARD; PRT; 274 AA.  
 AC P47678;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical lipoprotein MG440 precursor.  
 GN MG440.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RL Science 270:1397-403(1995).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
 CC (Potential).  
 CC -1- SIMILARITY: Belongs to the MG439 / MG440 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U39726; AAC72460.1; ALT\_INIT.  
 DR TIGR; MG440; --  
 DR InterPro; IPR001595; Lipoprotein\_3.  
 DR InterPro; IPR000437; Prok\_lipoprot\_5.  
 DR Pfam; PF00938; Lipoprotein\_3; 1.  
 DR ProDom; PD003276; Lipoprotein\_3; 1.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;  
 KW Complete proteome; Palmitate.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 274 HYPOTHETICAL LIPOPROTEIN MG440.  
 FT LIPID 26 26 N-palmitoyl cysteine (Potential).  
 FT LIPID 26 26 S-diacylglycerol cysteine (Potential).  
 SQ SEQUENCE 274 AA; 30777 MW; 6038B2290981F4C8 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```

RESULT 174
STR_STAAU STANDARD; PRT; 282 AA.
ID AC P12055;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE Streptomycin resistance protein.
GN STR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1280;
[1]
RN
RP SEQUENCE FROM N.A.
MEDLINE=88189810; PubMed=3357770;
RA Projan S.J., Moghazeh S., Novick R.P.;
RT "Nucleotide sequence of pS194, a streptomycin-resistance plasmid from
RT Staphylococcus aureus.";
RL Nucleic Acids Res. 16:2179-2187(1988).
CC -!- FUNCTION: THIS PROTEIN IS REQUIRED FOR STREPTOMYCIN RESISTANCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06627; CAA29839.1; -.
DR PIR; S00938; S00938.
DR InterPro; IPR007530; Adenyl_transf.
DR Pfam; PF04439; Adenyl_transf; 1.
DR Antibiotic resistance; Plasmid.
SQ SEQUENCE 282 AA; 33961 MW; 0C6A770634E5F58B CRC64;
Query Match 1.8%; Score 6; DB 1; Length 282;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 FNNILR 106
Db 182 FNNILR 187
|||||
-----
RESULT 175
YC80_GUITH STANDARD; PRT; 282 AA.
ID AC O78449;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Hypothetical 33.2 kDa protein ycf80.
GN YCF80.
OS Guillardia theta (Cryptomonas phi).
OC Chlorophyta.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
NCBI_TaxID=55529;
[1]
RN
RP SEQUENCE FROM N.A.
MEDLINE=99128221; PubMed=9929392;
RA Douglas S.E., Penny S.L.;
RT "The plastid genome of the cryptophyte alga, Guillardia theta:
RT complete sequence and conserved syntenic groups confirm its common
RT ancestry with red algae.";
RL J. Mol. Evol. 48:236-244(1999).
CC -!- SIMILARITY: Belongs to the ycf80 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

```



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF041468; AAC35638.1; -;  
 KW Hypothetical protein; Chloroplast.  
 SQ SEQUENCE 282 AA; 33221 MW; 0D84447DCADA943A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 82 LIADLQ 87  
 Db 210 LIADLQ 215  
 |||||

RESULT 176  
 AROE METH  
 ID AROE METH STANDARD; PRT; 283 AA.  
 AC 026344;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Shikimate 5-dehydrogenase (EC 1.1.1.25).  
 GN AROE OR MTH242.  
 OS Methanobacterium thermoautotrophicum.  
 CC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 CC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate +  
 CC NADPH.

CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC fourth step.  
 CC -!- SIMILARITY: Belongs to the shikimate dehydrogenase family.  
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A PROBABLE FRAMESHIFT  
 CC WAS CORRECTED AT POSITION 17.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE000811; AAB84748.1; ALT\_FRAME.  
 DR HAMAP; MF 00222; -; 1.  
 DR InterPro; IPR006152; Shikimate.  
 DR InterPro; IPR006151; Shikimate\_DH.  
 DR Pfam; PF01488; Shikimate\_DH; 1.  
 DR TIGRFAMs; TIGR00507; aroE; 1.  
 KW Aromatic amino acid biosynthesis; Oxidoreductase; NADP;  
 KW Complete proteome.  
 SQ SEQUENCE 283 AA; 30484 MW; 3293E04F92FC0B89 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 283;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 251 KPEMLK 256  
 Db 44 KPEMLK 49  
 |||||

RESULT 177  
 AROE CAUCR  
 ID AROE CAUCR STANDARD; PRT; 285 AA.  
 AC 09AC57;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Shikimate 5-dehydrogenase (EC 1.1.1.25).  
 GN AROE OR CC0003.  
 OS Caulobacter crescentus.  
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
 CC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Ely B.,  
 RA Poczka I., Nelson W.C., Newton A., Stephens C., Phacke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 CC -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate +  
 CC NADPH.

CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC fourth step.  
 CC -!- SIMILARITY: Belongs to the shikimate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE005675; AAK21991.1; -;  
 DR PIR; C87249; C87249.  
 DR TIGR; CC0003; -; 1.  
 DR HAMAP; MF 00222; -; 1.  
 DR InterPro; IPR006151; Shikimate\_DH.  
 DR Pfam; PF01488; Shikimate\_DH; 1.  
 KW Aromatic amino acid biosynthesis; Oxidoreductase; NADP;  
 KW Complete proteome.  
 SQ SEQUENCE 285 AA; 29014 MW; 91174F9C770C14FE CRC64;

Query Match 1.8%; Score 6; DB 1; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLLGE 234  
 Db 279 LKLLGE 284  
 |||||

RESULT 178  
 HUS1 SCHPO  
 ID HUS1 SCHPO STANDARD; PRT; 287 AA.  
 AC P78955;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein hus1.  
 GN HUS1 OR SPAC20G4.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 CC NCBI\_TaxID=4896;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 RP MEDLINE=97324587; PubMed=9180692;  
 RX Kostub C.F., Al-Khodairy F., Gharizadeh H., Carr A.M., Enoch T.;  
 RA "Molecular analysis of hus1+, a fission yeast gene required for S-M  
 RT and DNA damage checkpoints.";  
 RL Mol. Gen. Genet. 254:389-399(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Weijens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez C., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armetrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakowski G.V., Usery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Essential in controlling the S-M checkpoint that couples  
 CC mitosis to the completion of DNA replication. It is also required  
 CC for the response to DNA damage. Mutants defective in hus1 have a  
 CC nucleus that is cleaved by the septum or the septum divides the  
 CC cell into a nucleate and anucleate compartment.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y09438; CAA70588.1; -  
 CC DR EMBL; Z98600; CAB11254.1; -  
 CC DR PIR; T43396; T43396.  
 CC DR GeneDB-SPombe; SPAC20G4.04C; -  
 CC DR InterPro; IPR007150; Hus1.  
 CC DR Pfam; PF04005; Hus1; 1.  
 CC KW Mitosis; DNA damage; DNA repair.  
 CC SEQUENCE 287 AA; 32713 MW; 283CE8F538DFB78 CRC64;  
 Query Match 1.8%; Score 6; DB 1; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 208 TIFEDY 213

|||||  
 57 TIFEDY 62  
 Db  
 RESULT 179  
 SUCD THETH  
 ID SUCD THETH STANDARD; PRT; 288 AA.  
 AC P09143;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Succinyl-CoA synthetase alpha chain (EC 6.2.1.5) (SCS-alpha).  
 GN SUCD OR SCSA.  
 OS Thermus thermophilus.  
 CC Bacteria; Deinococcus-Thermus; Deinococci; Thermaceae;  
 CC Thermus.  
 CC NCBI\_TaxID=274;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 RP STRAIN=B / NCIB 11247;  
 RX MEDLINE=89041573; PubMed=3186449;  
 RA Nicholls D.J., Sundaram T.K., Atkinson T., Minton N.P.;  
 RT "Nucleotide sequence of the succinyl-CoA synthetase alpha-subunit  
 RL from Thermus aquaticus B.";  
 RN Nucleic Acids Res. 16:9858-9858(1988).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 RP STRAIN=B / NCIB 11247;  
 RX MEDLINE=90375010; PubMed=2204576;  
 RA Nicholls D.J., Sundaram T.K., Atkinson T., Minton N.P.;  
 RT "Cloning and nucleotide sequences of the mdh and sucD genes from  
 RL Thermus aquaticus B.";  
 RN FEMS Microbiol. Lett. 58:7-14(1990).  
 CC [3]  
 CC SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 33923 / DSM 674 / AT-62;  
 RX MEDLINE=91238680; PubMed=2034208;  
 RA Nishiyama M., Horinouchi S., Bepu T.;  
 RT "Characterization of an operon encoding succinyl-CoA synthetase and  
 RL malate dehydrogenase from Thermus flavus AT-62 and its expression in  
 RL Escherichia coli.";  
 RL Mol. Gen. Genet. 226:1-9(1991).  
 CC -!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +  
 CC phosphate.  
 CC -!- PATHWAY: Tricarboxylic acid cycle.  
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains.  
 CC -!- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS ATP AND CATALYZES  
 CC PHOSPHORYL TRANSFER TO ONE OF ITS HISTIDINE RESIDUES. THE COMPLETE  
 CC ACTIVE SITE IS PROBABLY LOCATED IN THE REGION OF ALPHA-BETA  
 CC CONTACT.  
 CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase alpha  
 CC subunit family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M35832; AAA27504.2; -  
 CC DR EMBL; X56033; CAA39507.1; -  
 CC DR EMBL; X54073; CAA38007.1; -  
 CC DR HSP; P07459; ISCU.  
 CC DR InterPro; IPR003781; CoA binding.  
 CC DR InterPro; IPR005810; CoA ligase.  
 CC DR InterPro; IPR005811; CoA ligase.  
 CC DR Pfam; PF02629; CoA binding; 1.  
 CC DR Pfam; PF00549; ligase-CoA; 1.  
 CC DR PRINTS; PR01798; SCSASYNTHASE.  
 CC DR TIGRFAMS; TIGR01019; succoalpha; 1.  
 CC DR PROSITE; PS01216; SUCCINYL\_COA\_LIG\_1; 1.





TIGR; SPI212; -.  
 DR HAMAP; MF 01080; -. 1;  
 DR InterPro; IPR004510; TruB.  
 DR InterPro; IPR002501; TruB\_synth\_N.  
 DR Pfam; PF01509; TruB\_N; 1.  
 DR TIGRFAMs; TIGR00431; TruB; 1.  
 KW tRNA processing; Lyase; Complete proteome.  
 FT ACT SITE 38 BY SIMILARITY.  
 SQ SEQUENCE 292 AA; 32271 MW; B7A192469C4BC168 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LAILEX 30  
 Db 275 LAILEX 280  
 |||||

RESULT 184  
 TRUB\_STR6  
 ID TRUB\_STR6 STANDARD; PRT; 292 AA.  
 AC Q8CWR2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE tRNA pseudouridine synthase B (EC 4.2.1.70) (tRNA pseudouridine 55  
 DE synthase) (Pseudouridylyl synthase) (Uracil  
 DE hydrolyase).  
 GN TRUB OR SPRI092.  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 CC Streptococcus.  
 OX NCBI\_TaxID=171101;  
 RN [1]\_TaxID=171101;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilmore R., Glass J.S., Kioja H., Kraft A.R., Lagace R.E.,  
 RA Lelanc D.-J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,  
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 CC -!- FUNCTION: Responsible for synthesis of pseudouridine from  
 CC uracil-55 in the psi GC loop of transfer RNAs (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine  
 CC 5'-phosphate + H(2)O.  
 CC -!- SIMILARITY: Belongs to the pseudouridine synthase truB family.  
 CC Subfamily 1.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE008482; AAK99895.1; -.  
 DR HAMAP; MF 01080; -. 1.  
 DR InterPro; IPR004510; TruB.  
 DR InterPro; IPR002501; TruB\_synth\_N.  
 DR Pfam; PF01509; TruB\_N; 1.  
 DR TIGRFAMs; TIGR00431; TruB; 1.  
 KW tRNA processing; Lyase; Complete proteome.  
 FT ACT SITE 38 BY SIMILARITY.  
 SQ SEQUENCE 292 AA; 32256 MW; 2716D45FFC2E145E CRC64;

Query Match 1.8%; Score 6; DB 1; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LAILEX 30  
 Db 275 LAILEX 280  
 |||||

RESULT 185  
 SUO1\_RAT  
 ID SUO1\_RAT STANDARD; PRT; 295 AA.  
 AC P52844;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Estrogen sulfotransferase, isoform 1 (EC 2.8.2.4) (EST-1)  
 DE (Sulfotransferase, estrogen-prefering) (Estrone sulfotransferase).  
 GN STE1 OR SULT1E1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fischer 344; TISSUE=Liver;  
 RX MEDLINE=96305357; PubMed=8688469;  
 RA Rikke B.A., Roy A.K.;  
 RT "Structural relationships among members of the mammalian  
 RT sulfotransferase gene family.";  
 RL Biochim. Biophys. Acta 1307:331-338(1996).  
 CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the  
 CC level of the estrogen receptor by sulfonylating free estradiol (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + estrone =  
 CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U50204; AAB07680.1; -.  
 DR HSP; P49891; 1AQU  
 DR InterPro; IPR000863; Sulfotransferase.  
 DR Pfam; PF00685; Sulfotransfer; 1.  
 DR ProDom; PD001218; Sulfotransferase; 1.  
 KW Transferase; Steroid-binding.  
 FT BINDING 259 265  
 SQ SEQUENCE 295 AA; 35509 MW; 696A12FDA923A12E CRC64;

Query Match 1.8%; Score 6; DB 1; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIEFL 305  
 Db 206 KLIEFL 211  
 |||||

RESULT 186  
 SUO2\_RAT  
 ID SUO2\_RAT STANDARD; PRT; 295 AA.  
 AC P52845;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

30-MAY-2000 (Rel. 39, Last annotation update)  
DE Estrogen sulfotransferase, isoform 2 (EC 2.8.2.4) (EST-2)  
DE (Sulfotransferase, estrogen-preferring) (Estrone sulfotransferase).  
GN STP2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Fischer 344; TISSUE=Liver;  
RC MEDLINE=96305357; PubMed=8688469;  
RA Rikke B.A., Roy A.K.;  
RX "Structural relationships among members of the mammalian  
RT sulfotransferase gene family";  
RL Biochim. Biophys. Acta 1307:331-338(1996).  
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the  
CC level of the estrogen receptor by sulfonylating free estradiol.  
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + estrone =  
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U50205; AAB07681.1; -.  
CC HSSP; P49891; 1AQU.  
CC InterPro; IPR000863; Sulfotransferase.  
CC Pfam; PF00685; Sulfotransfer; 1.  
CC ProDom; PD001218; Sulfotransferase; 1.  
CC TrnTransferase; Steroid-binding.  
FT BINDING 259 265 PAPS BINDING SITE (POTENTIAL).  
SQ SEQUENCE 295 AA; 35364 MW; 149B5C9D46039AAF CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 300 KLIEFL 305  
Db 206 KLIEFL 211  
|||||  
-----  
RESULT 187  
SUO3 RAT STANDARD; PRT; 295 AA.  
ID SUO3 RAT  
AC P49889,  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Estrogen sulfotransferase, isoform 3 (EC 2.8.2.4) (EST-3)  
DE (Sulfotransferase, estrogen-preferring) (Estrone sulfotransferase).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 258-265.  
RP TISSUE=Liver;  
RC MEDLINE=92261615; PubMed=1374839;  
RA Demian W.F., Song C.S., Kim D.S., Her S., Gallwitz W., Rao T.R.,  
RA Slowczynska M., Chatterjee B., Roy A.K.;  
RX "Estrogen sulfotransferase of the rat liver: complementary DNA  
RT cloning and age- and sex-specific regulation of messenger RNA";  
RL Mol. Endocrinol. 6:589-597(1992).

[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Liver;  
RC MEDLINE=95161323; PubMed=7857871;  
RA Falany J.L., Krasnykh V., Mikhieva G., Falany C.N.;  
RX "Isolation and expression of an isoform of rat estrogen  
RT sulfotransferase";  
RL J. Steroid Biochem. Mol. Biol. 52:35-44(1995).  
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the  
CC level of the estrogen receptor by sulfonylating free estradiol.  
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + estrone =  
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- TISSUE SPECIFICITY: LIVER OF YOUNG MATURE MALES AND UTERUS.  
CC -!- DEVELOPMENTAL STAGE: Expressed only in the liver of young adult  
CC animals (100 days old) and is absent in the prepubertal male (27  
CC days old), senescent male (800 days old) and female liver.  
CC -!- INDUCTION: Induced by androgens and suppressed by estrogens. The  
CC expression is under the influence of pituitary growth hormone and  
CC thyroid hormone. Is regulated by progesterone in the uterus.  
CC -!- SIMILARITY: Belongs to the sulfotransferase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M86758; AAA41128.1; -.  
CC EMBL; S76489; AAB33441.1; -.  
CC PIR; A41930; A41930.  
CC HSSP; P49891; 1AQU.  
CC InterPro; IPR000863; Sulfotransferase.  
CC Pfam; PF00685; Sulfotransfer; 1.  
CC ProDom; PD001218; Sulfotransferase; 1.  
CC TrnTransferase; Steroid-binding; Multigene family.  
FT BINDING 259 265 PAPS BINDING SITE (POTENTIAL).  
FT CONFLICT 150 150 P -> Q (IN REF. 2).  
FT CONFLICT 238 238 T -> I (IN REF. 2).  
FT CONFLICT 295 295 L -> P (IN REF. 2).  
SQ SEQUENCE 295 AA; 35415 MW; A77807A21DD2E7EB CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 300 KLIEFL 305  
Db 206 KLIEFL 211  
|||||  
-----  
RESULT 188  
SUO6 RAT STANDARD; PRT; 295 AA.  
ID SUO6 RAT  
AC P49830,  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Estrogen sulfotransferase, isoform 6 (EC 2.8.2.4) (EST-6)  
DE (Sulfotransferase, estrogen-preferring) (Estrone sulfotransferase).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Liver;  
RC MEDLINE=95161323; PubMed=7857871;  
RA Falany J.L., Krasnykh V., Mikhieva G., Falany C.N.;  
RX "Isolation and expression of an isoform of rat estrogen



```
RT sulfotransferase.";
RL J. Steroid Biochem. Mol. Biol. 52:35-44(1995).
CC -!- FUNCTION: Sulfation of estrone and estradiol. May control the
CC level of the estrogen receptor by sulfonylating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + estrone =
CC adenosine 3',5'-bisphosphate + estrone 3-sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DEVELOPMENTAL STAGE: Expressed in males rats, but not in females.
CC -!- INDUCTION: Induced by androgens and suppressed by estrogens.
CC Expression is under the influence of pituitary growth hormone and
CC thyroid hormone.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S76490; AAB33442.1; -.
DR HSRF; P49891; IAOU.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase; Steroid-binding; Multigene family.
FT BINDING 259 265 PAPS BINDING SITE (POTENTIAL).
SQ SEQUENCE 295 AA; 35302 MW; 9FED5861AFDC9B05 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 295;
Best Local Similarity 100.0%; Pred.No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 300 KLIIEFL 305
Db 206 KLIIEFL 211
|||||

RESULT 189
SUOT_MOUSE
ID SUOT_MOUSE STANDARD; PRT; 295 AA.
AC P49891;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Estrogen sulfotransferase, testis isoform (EC 2.8.2.4)
DE (Sulfotransferase, estrogen-prefering).
DE STE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=95269690; PubMed=7750469;
RA Song W.-C., Moore R., McLachlan J.A., Negishi M.;
RT "Molecular characterization of a testis-specific estrogen
RT sulfotransferase and aberrant liver expression in obese and
RT diabetic C57BL/KSJ-db/db mice.";
RL Endocrinology 136:2477-2484(1995).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS), AND REVISION TO 113.
RC STRAIN=C57BL/KSJ-DB/DB; TISSUE=Testis;
RX MEDLINE=98020506; PubMed=9360604;
RA Kakuta Y., Pedersen L.G., Carter C.W., Negishi M., Pedersen L.C.;
RT "Crystal structure of estrogen sulfotransferase.";
RL Nat. Struct. Biol. 4:904-908(1997).
CC -!- FUNCTION: Sulfation of estradiol and estrone. May control the
CC level of the estrogen receptor by sulfonylating free estradiol.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + estrone =
```



```

FT STRAND 248 248
FT TURN 250 252
FT HELIX 263 266
FT TURN 267 267
FT HELIX 270 284
FT TURN 285 286
SQ SEQUENCE 295 AA; 35590 MW; 8E85AB47952BFB1C CRC64;

Query Match 1.8%; Score 6; DB 1; Length 295;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIEFL 305
Db 206 KLIEFL 211

RESULT 190
XERC_PASMU
ID_XERC_PASMU STANDARD; PRT; 295 AA.
AC Q9CKC2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Tyrosine recombinase xerc.
GN XERC OR PM1701.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: Site-specific tyrosine recombinase, which acts by
CC catalyzing the cutting and rejoining of the recombining DNA
CC molecules. The xerc-xerc complex is essential to convert dimers of
CC the bacterial chromosome into monomers to permit their segregation
CC at cell division. It also contributes to the segregational
CC stability of plasmids (By similarity).
CC -!- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two
CC molecules of xerc and two molecules of xerD (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the "phage" integrase family. Xerc
CC subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006206; AAK03785.1; -.
CC HSP; F21891; LAOP.
CC HAMAP; MF_01808; -.
CC InterPro; IPR004107; Phage integr. N.
CC InterPro; IPR002104; Phage integrase.
CC Pfam; PF02899; Phage_integr. N; 1.
CC Pfam; PF00589; Phage_integrase; 1.
CC DNA recombination; DNA integration; Cell division;
CC Chromosome partition; DNA-binding; Complete proteome.
CC ACT_SITE 145 145
CC ACT_SITE 169 169
CC ACT_SITE 237 237
CC ACT_SITE 240 240
CC ACT_SITE 263 263
CC ACT_SITE 272 272
CC TRANSIENT COVALENT LINKAGE TO DNA DURING
CC STRAND CLEAVAGE AND REJOINING (BY

FT SEQUENCE 295 AA; 34206 MW; 99FD70DF1E37BA1D CRC64;
SIMILARITY).
SQ
Query Match 1.8%; Score 6; DB 1; Length 295;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGL 78
Db 139 LYSSGL 144

RESULT 191
ERA_FUSNN
ID_ERA_FUSNN STANDARD; PRT; 296 AA.
AC Q8RGM1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE GTP-binding protein era homolog.
GN ERA OR FN0270.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]_
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyripides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
CC -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
CC and is essential for cell growth (By similarity).
CC -!- SIMILARITY: Belongs to the era/trmG family of GTP-binding
CC proteins. Era subfamily.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE010538; AAJ94476.1; -.
CC HAMAP; MF_00367; -.
CC InterPro; IPR005662; Era.
CC InterPro; IPR005289; GTP-binding_dom.
CC InterPro; IPR006073; GTP1_OBG.
CC InterPro; IPR004087; KH_dom.
CC InterPro; IPR009019; KH_prok.
CC InterPro; IPR004044; KH_TYPE_2.
CC InterPro; IPR002917; MMR_HSR1.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00013; KH; 1.
CC Pfam; PF01926; MMR_HSR1; 1.
CC PRINTS; PR00326; GTP1_OBG.
CC TIGRFAMs; TIGR00436; era; 1.
CC TIGRFAMs; TIGR00650; MG442; 1.
CC TIGRFAMs; TIGR00231; small_GTP; 1.
CC PROSITE; PS50823; KH_TYPE_2; 1.
CC GTP-binding; RNA-binding; Complete proteome.
CC NP_BIND 10 17
CC NP_BIND 57 61
CC NP_BIND 120 123
CC NP_BIND 202 279
CC DOMAIN

```

```

SQ SEQUENCE 296 AA; 33929 MW; B05FPB966695E3969 CRC64;
  Query Match 1.8%; Score 6; DB 1; Length 296;
  Best Local Similarity 100.0%; Pred. No. 3.3e+02;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 VKLKD 23
Db 77 VKLKD 82

RESULT 192
ID TSY AGABI STANDARD; PRT; 296 AA.
AC Q9P4T7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thymidylate synthase (EC 2.1.1.45) (TS) (TSase).
GN TMS1.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OK NCBI_TaxID=5341;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hors1 U3;
RA Eastwood D.C., Bains N.K., Henderson J., Burton K.S.;
RT "Genomic sequencing of superoxide dismutase in Agaricus bisporus.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
CC dihydrofolate + dTMP.
CC -1- PATHWAY: Deoxyribonucleotide biosynthesis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: Belongs to the thymidylate synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ401221; CAB96042.1; -.
CC HSSP; P04819; 1HW4.
CC InterPro; IPR000398; Thymidylat synth.
CC Pfam; PF00303; thymidylat synt; 1.
CC PRINTS; PR00108; THYMDSNTASE.
CC ProDom; PD001180; Thymidylat synt; 1.
CC ProDom; PD001180; Thymidylat synt; 1.
CC PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
KW Transferase, Methyltransferase; Nucleotide biosynthesis.
FT ACT_SITE 171 171 BY SIMILARITY.
SQ SEQUENCE 296 AA; 33679 MW; 8F8A8989D4A952A CRC64;

  Query Match 1.8%; Score 6; DB 1; Length 296;
  Best Local Similarity 100.0%; Pred. No. 3.3e+02;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 LRECIR 151
Db 137 LRECIR 142

RESULT 193
ID HEM3 FUSNN STANDARD; PRT; 298 AA.
AC Q8RFP5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
DE synthase) (HMBs) (Pre-uroporphyrinogen synthase).

```

---

```

GN HEMC OR FN0645.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
EX MEDLINE=21886394; PubMed=1189109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kypides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
CC hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
CC hydroxymethylbilane + 4 NH(3).
CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
CC porphobilinogen subunits are added (By similarity).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the HMBs family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF010575; AAL94841.1; -.
CC HAMAP; MF 00260; -.
CC InterPro; IPR000860; Porphobil deam.
CC Pfam; PF01379; Porphobil deam; 1.
CC Pfam; PF03900; Porphobil deamC; 1.
CC PRINTS; PR00151; PORPHBDMNASE.
CC ProDom; PD002745; Porphobil deam; 1.
CC ProDom; PD002745; Porphobil deam; 1.
CC TIGRFAMs; TIGR00212; hemC; 1.
CC DR PROSITE; PS00533; PORPHOBILINOGEN_DEAM; 1.
KW Porphyrin biosynthesis; Transferase; Complete proteome.
FT BINDING 242 242 PYRROMETHANE COFACTOR (BY SIMILARITY).
SQ SEQUENCE 298 AA; 33098 MW; 13908A7D0AA56984 CRC64;

  Query Match 1.8%; Score 6; DB 1; Length 298;
  Best Local Similarity 100.0%; Pred. No. 3.3e+02;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 77 GLLVTL 82
Db 114 GLLVTL 119

RESULT 194
ID RT03 ACACA STANDARD; PRT; 298 AA.
AC P46754;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Mitochondrial ribosomal protein S3.
GN RPS3.
OS Acanthamoeba castellanii (Amoeba).
OC Mitochondrion.
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OK NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30010 / Neff;

```

RA MEDLINE=95147275; PubMed=7844823;  
RA Burger G., Plante I., Lonergan K.M., Gray M.W.;  
RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba  
RT castellanii: complete sequence, gene content and genome  
RL J. Mol. Biol. 245:522-537(1995).  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- SIMILARITY: Belongs to the S3P family of ribosomal proteins.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U12386; RAD11841.1; -.  
DR PIR; S53849; S53849.  
DR InterPro; IPR009019; KH\_prok.  
DR InterPro; IPR001351; Ribosomal\_S3\_C.  
DR InterPro; IPR008282; Ribosomal\_S3\_N.  
DR Pfam; PF00189; Ribosomal\_S3\_C; 1.  
DR Pfam; PF00417; Ribosomal\_S3\_N; 1.  
DR PROSITE; PS00548; RIBOSOMAL\_S3; FALSE\_NEG.  
KW Ribosomal protein; Mitochondrion.  
SQ SEQUENCE 298 AA; 36060 MW; 29415935EE187DE6 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 126 LFMLLK 131  
| | | | |  
DB 125 LFMLLK 130  
  
RESULT 195  
XERD\_LEPIN  
ID XERD\_LEPIN STANDARD; PRT; 298 AA.  
AC Q7ZAN7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE Tyrosine recombinase xerD.  
GN XERD OR LA2483.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN [1]\_TaxID=173;  
RP SEQUENCE FROM N.A.  
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
RX MEDLINE=22598143; PubMed=12712204;  
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,  
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,  
RA Jiang Y.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,  
RA Zhang Y., Zhu G.-F., Wang B.-Q., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
RA Yao Z.-J., Shen Y., Qian B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,  
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
RA Xu J.-G., Zhao G.-P.;  
RT "Unique physiological and pathogenic features of Leptospira  
RT interrogans revealed by whole-genome sequencing."  
RL Nature 422:888-893(2003).  
CC -!- FUNCTION: Site-specific tyrosine recombinase, which acts by  
CC catalyzing the cutting and rejoining of the recombining DNA  
CC molecules. The xerC-xerD complex is essential to convert dimers of  
CC the bacterial chromosome into monomers to permit their segregation  
CC at cell division. It also contributes to the segregational  
CC stability of plasmids (By similarity).  
CC -!- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two  
CC molecules of xerC and two molecules of xerD (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the "phage" integrase family. XerD

CC subfamily 1.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE011415; AAN49682.1; -.  
DR HAMAP; MF 01807; -; 1.  
DR InterPro; IPR004107; Phage\_integr\_N.  
DR InterPro; IPR002104; Phage\_integrase.  
DR Pfam; PF02899; Phage\_integr\_N; 1.  
DR Pfam; PF00589; Phage\_integrase; 1.  
KW DNA recombination; DNA integration; Cell division;  
KW Chromosome partition; DNA-binding; Complete proteome.  
FT ACT\_SITE 149 149 BY SIMILARITY.  
FT ACT\_SITE 173 173 BY SIMILARITY.  
FT ACT\_SITE 244 244 BY SIMILARITY.  
FT ACT\_SITE 247 247 BY SIMILARITY.  
FT ACT\_SITE 270 270 BY SIMILARITY.  
FT ACT\_SITE 279 279 BY SIMILARITY.  
FT TRANSIENT COVALENT LINKAGE TO DNA DURING  
FT STRAND CLEAVAGE AND REJOINING (BY  
FT SIMILARITY).  
SQ SEQUENCE 298 AA; 35073 MW; F1578AA6F23FEB25 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 298;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 73 LYSSGL 78  
| | | | |  
DB 143 LYSSGL 148  
  
RESULT 196  
G3P\_DICDI  
ID G3P\_DICDI STANDARD; PRT; 299 AA.  
AC Q94469;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-NAR-2004 (Rel. 43, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)  
DE (fragment).  
GN GPDA.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]\_TaxID=44689;  
RP SEQUENCE FROM N.A.  
RC STRAIN=AX4;  
RX MEDLINE=97130906; PubMed=8976605;  
RA Roger A.J., Smith M.W., Doolittle R.F., Doolittle W.F.;  
RT "Evidence for the Heterobosea from phylogenetic analysis of genes  
RT encoding glyceraldehyde-3-phosphate dehydrogenase."  
RL J. Eukaryot. Microbiol. 43:475-485(1996).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -!- PATHWAY: Second phase of glycolysis; first step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC      EMBL; U55243; AAC47285.1; -.
DR      HSSP; P56649; IS2J.
DR      DictyBase; DB0185087; gpdA.
DR      InterPro; IPR000173; GAP_dehydrogenase.
DR      InterPro; IPR006424; GAPDH-I.
DR      Pfam; PF00044; gpdh; 1.
DR      Pfam; PF02800; gpdh; C; 1.
DR      PRINTS; PR00078; G3PDHGRNASE.
DR      TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR      PROSITE; PS00071; GAPDH; 1.
DR      Glycolysis; Oxidoreductase; NAD.
FT      NON TER 1
FT      BINDING 138 138
FT      ACT_SITE 165 165
FT      NON TER 299 299
FT      SEQUENCE 299 AA; 32407 MW; B5948FA38F8606F5 CRC64;
Query Match      1.8%; Score 6; DB 1; Length 299;
Best Local Similarity 100.0%; Pred No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      154 PLAKII 159
DB      145 PLAKII 150
|||||
RESULT 197
HIS1_BUCBP
ID      HIS1_BUCBP      STANDARD;      PRT;      299 AA.
AC      P59453;
DT      10-OCT-2003 (Rel. 42, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      ATP phosphoribosyltransferase (EC 2.4.2.17).
DE      HSG OR BHP093.
OS      Buchnera aphidicola (subsp. Baizongia pistaciae).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Buchnera.
OC      NCBI_TaxID=135842;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22426901; PubMed=12522265;
RX      Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA      Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA      Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT      "Reductive genome evolution in Buchnera aphidicola.";
RL      Proc. Natl. Acad. Sci. U.S.A. 100:581-586 (2003).
CC      -!- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribosyl)-ATP + diphosphate =
CC      ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
CC      -!- PATHWAY: Histidine biosynthesis; first step. Very important in the
CC      regulation of histidine metabolism.
CC      -!- SUBUNIT: Homohexamer (By similarity).
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -!- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.
CC      Long subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE014016; AA026828.1; -.
DR      HAMAP; MF_00079; -.
DR      InterPro; IPR001348; ATP_phospho_trans.
DR      Pfam; PF01634; Hsg; 1.
DR      TIGRFAMs; TIGR00070; hsg; 1.
DR      PROSITE; PS01316; ATP_P_PHORIBOSYLTR; 1.

```

```
RESULT 199
FTRL METKA
ID FTRL METKA STANDARD; PRT; 300 AA.
AC Q8TX60;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formylmethanofuran--tetrahydromethanopterin formyltransferase-like
DE protein.
GN MK0816.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and morphology of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -!- SIMILARITY: Belongs to the FTR family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AEO10372; AAO02029.1; -.
CC HAMAP; MF 00579; atypical; 1.
CC InterPro; IPR002770; FTR.
CC Pfam; PF01913; FTR; 1.
CC Pfam; PF02741; FTR_C; 1.
CC PIRSF; PIRSF006414; FTR; 1.
CC ProDom; PD007702; FTR; 1.
CC Transferrase; Complete proteome.
SQ SEQUENCE 300 AA; 32966 MW; EF08FC68601ADCOE CRC64;

Query Match 1.8%; Score 6; DB 1; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 LRECIR 151
DB 294 LRECIR 299
|||||

RESULT 200
SC14_KLUJA
ID SC14_KLUJA STANDARD; PRT; 301 AA.
AC P24859;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SEC14 cytosolic factor (Phosphatidylinositol/phosphatidylcholine
DE transfer protein) (PI/PC TP).
GN SEC14.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90330560; PubMed=2198263;
RA Salama S.R., Cleves A.E., Malehorn D.E., Whitters E.A.,
RA Bankaitis V.A.;

Query Match 1.8%; Score 6; DB 1; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 LRECIR 151
DB 294 LRECIR 299
|||||

RESULT 201
YF85_MYCPN
ID YF85_MYCPN STANDARD; PRT; 302 AA.
AC P75195;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical lipoprotein MPN585 precursor (D02_orf302).
GN MPN585 OR MP257.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE MG067 / MG068 / MG395 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000025; AAB95905.1; -.
CC PIR; S73583; S73583.
CC InterPro; IPR002414; DUF30/31.
CC InterPro; IPR000437; Prok_lipoprot_s.
CC Pfam; PF01732; DUF31; 1.
CC PRINTS; PR00840; Y06768FAMILY.
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.

RT "Cloning and characterization of Kluyveromyces lactis SEC14, a gene
RT whose product stimulates Golgi secretory function in Saccharomyces
RT cerevisiae."
RL J. Bacteriol. 172:4510-4521(1990).
CC -!- FUNCTION: Required for transport of secretory proteins from the
CC Golgi complex. Catalyzes the transfer of phosphatidylinositol and
CC phosphatidylcholine between membranes in vitro (By similarity).
CC -!- SUBCELLULAR LOCATION: Associated with the Golgi complex as a
CC peripheral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 CRAL-TRIO domain.
DR PIR; A37766; A37766.
DR HSSP; P24280; 1AUA.
DR InterPro; IPR001251; CRAL_TRIO_C.
DR InterPro; IPR008273; CRAL_TRIO_N.
DR InterPro; IPR001071; RetBind/tocTrans.
DR Pfam; PF00650; CRAL_TRIO; 1.
DR Pfam; PF03765; CRAL_TRIO_N; 1.
DR PRINTS; PR00180; CRETINALDHPP.
DR SMART; SM00516; SEC14; 1.
DR PROSITE; PS0191; CRAL_TRIO; 1.
KW Transport; Protein transport; Golgi stack.
FT DOMAIN 97 270
FT SEQUENCE 301 AA; 34529 MW; 8B8353752FF5FCE4 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 DTIFED 212
DB 87 DTIFED 92
|||||
```

```

KW Hypothetical protein; Lipoprotein; Membrane; Signal;
KW Complete proteome; Palmitate.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 302 HYPOTHETICAL LIPOPROTEIN MPN585.
FT LIPID 23 23 N-palmitoyl cysteine (Potential).
FT LIPID 23 23 S-diacylglycerol cysteine (Potential).
SQ SEQUENCE 302 AA; 35102 MW; 0D2B90B2DC7FC78A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PLSKS 10
Db 213 PLSKS 218

RESULT 202
Y191_METTH STANDARD; PRT; 304 AA.
AC 026293;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamine amidotransferase-like protein MTH191.
GN MTH191.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000806; AAB84697.1; -
CC PIR; A69122; A69122.
CC HSSP; P00497; 1A00.
CC InterPro; IPR000583; GATase_2.
CC Pfam; PF00310; GATase_2; 1.
CC PROSITE; PS00443; GATASE_TYPE_II; FALSE_NEG.
KW Transferase; Glutamine amidotransferase; Complete proteome.
FT INT MET 0 0 BY SIMILARITY.
FT ACT SITE 1 1 GATASE (BY SIMILARITY).
SQ SEQUENCE 304 AA; 33779 MW; 02F652DE0FF6FC89 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 ASEVS 43
Db 276 ASEVS 281

RESULT 203
NUIM_APIIL STANDARD; PRT; 305 AA.
AC P3487;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
GN ND1.
OS Apis mellifera ligustica (Common honeybee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Megaloptera; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7469;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thorax;
RX MEDLINE=93114603; PubMed=8417993;
RA Crozier R.H., Crozier Y.C.;
RT "The mitochondrial genome of the honeybee Apis mellifera: complete
RT sequence and genome organization.";
RL Genetics 133:97-117(1993).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SIMILARITY: Belongs to the complex I subunit 1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L06178; AAB96810.1; -
CC PIR; S52972; S52972.
CC InterPro; IPR001694; Resp_NADH_dhl.
CC Pfam; PF00146; NADHdh; 1.
CC PROSITE; PS00667; COMPLEX1_ND1_1; FALSE_NEG.
CC PROSITE; PS00668; COMPLEX1_ND1_2; FALSE_NEG.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 305 AA; 36838 MW; 01C4F46DFCF7A6E6 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 305;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 ILFML 130
Db 101 ILFML 106

RESULT 204
PPE1_SCHPO STANDARD; PRT; 305 AA.
AC P36614;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein phosphatase ppe1 (EC 3.1.3.16) (Phosphatase
DE epl).
GN PPE1 OR EPL1 OR PPX1 OR SPC1739.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972 / HM123;
RX MEDLINE=93250325; PubMed=8387356;
RA Shimanuki M., Kinoshita N., Ohkura H., Yoshida T., Toda T.,

```



RA Yanagida M.;  
RT "Isolation and characterization of the fission yeast protein  
RT phosphatase gene ppe1+ involved in cell shape control and mitosis.";  
RL Mol. Biol. Cell 4:303-313(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93250328; PubMed=8387358;  
RA Matsumoto T., Beach D.;  
RA "Interaction of the pml1/spi1 mitotic checkpoint with a protein  
RT phosphatase.";  
RL Mol. Biol. Cell 4:337-345(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Oliver K., O'Neil S., Saunders D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Quail M.A., Rabinowitsch E.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymouprez B.,  
RA Weijens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Bozycz K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Gadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
RN [4]  
RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.  
RX MEDLINE=22658053; PubMed=12773390;  
RA Goshima G., Iwasaki O., Obuse C., Yanagida M.;  
RT "The role of ppe1/ppp phosphatase for equal chromosome segregation in  
RT fission yeast kinetochore.";  
RL EMBO J. 22:2752-2763(2003).  
CC -!- FUNCTION: Has a role in chromosome segregation. May provide a  
CC dynamic connection between kinetochore microtubules and  
CC kinetochore chromatin. Negatively regulates mis12.  
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
CC phosphate.  
CC -!- COFACTOR: Binds 1 iron ion and 1 manganese ion per subunit (By  
CC similarity).  
CC -!- SUBUNIT: Interacts with stp5, ekl1 and mis12.  
CC -!- SUBCELLULAR LOCATION: Nuclear; associated with chromatin.  
CC -!- SIMILARITY: Belongs to the ppp phosphatase family. PP-V subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: D13712; BAA02865.1; --  
CC EMBL: Z18925; CAA79358.1; --  
CC EMBL: AL031540; CAA20786.1; --  
CC PIR: A47727; A47727.  
CC HSSP: P08129; 1FUJ.  
CC GeneDB\_Spombe; SPCC1739.12; --  
DR GO: 0007049; P: cell cycle; ISS.  
DR GO: 0016043; P: cell organization and biogenesis; ISS.  
DR InterPro: IPR004843; M: ppe1.  
DR Pfam: PF001149; Metallophos; 1.  
DR PRINTS: PR00114; STPHPTASE.  
DR ProDom: PD000252; T: phtase\_apah; 1.  
DR SMART: SM00156; PP2AC; 1.  
DR PROSITE: PS00125; SER, THR, PHOSPHATASE; 1.  
KW Hydroxylase; Metal-binding; Iron; Manganese; Cell cycle; Mitosis;  
KW Nuclear protein.  
FT METAL 51 51 IRON (BY SIMILARITY).  
FT METAL 53 53 IRON AND MANGANESE (BY SIMILARITY).  
FT METAL 79 79 MANGANESE (BY SIMILARITY).  
FT METAL 111 111 GENERAL ACID (BY SIMILARITY).  
FT ACT\_SITE 112 112 MANGANESE (BY SIMILARITY).  
FT METAL 161 161 MANGANESE (BY SIMILARITY).  
FT METAL 235 235 MANGANESE (BY SIMILARITY).  
SQ SEQUENCE 305 AA; 35259 MW; 001980A1CC7646D7 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 305;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;  
QY 126 LFMLLK 131  
DB 93 LFMLLK 98  
RESULT 205  
TRUB HELMO STANDARD; PRT; 307 AA.  
ID Q8GB93; (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, last sequence update)  
DT 10-OCT-2003 (Rel. 42, last annotation update)  
DE tRNA pseudouridine synthase B (EC 4.2.1.70) (tRNA pseudouridine 55  
DE synthase) (Psi55 synthase) (Pseudouridylate synthase) (Uracil  
DE hydrolase) (Fragment).  
DE TRUB.  
GN Helicobacillus mobilis.  
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;  
OC Helicobacillus.  
OX NCBI\_TaxID=28064;  
RN [1] SEQUENCE FROM N.A.  
RC STRAIN=RHMO02226;  
RA Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,  
RA Gerdess S., Kyrtides N., Overbeek R.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Responsible for synthesis of pseudouridine from  
CC uracil-55 in the psi GC loop of transfer RNAs (By similarity).  
CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine  
CC 5'-phosphate + H(2)O.  
CC -!- SIMILARITY: Belongs to the pseudouridine synthase trub family.  
CC Subfamily 1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: AY142837; AAN87441.1; --  
CC HAMAP: MF 01080; --; 1.  
DR InterPro: IPR004510; Trub.  
DR InterPro: IPR002501; Trub\_synth\_N.  
DR Pfam: PF01509; Trub\_N, 1.  
DR TIGRFAMs: TIGR00431; Trub; 1.  
KW tRNA processing; Lyase.  
FT ACT\_SITE 45 45 BY SIMILARITY.



```

FT NON TER 307 307
SQ SEQUENCE 307 AA; 34836 MW; 9AD3D85DAF098EEE CRC64;

Query Match 1.8%; Score 6; DB 1; Length 307;
Best Local Similarity 100.0%; Pred.No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 GELIID 238
DB 166 GELIID 171

RESULT 206
MRAW_HELPY
ID MRAW_HELPY STANDARD; PRT; 308 AA.
AC O25411;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR HP0707.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Karlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khakhria H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RA "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (by similarity).
CC -!- SIMILARITY: Belongs to the mraw family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE000584; AAD07761.1; -.
CC PIR; C64608; C64608.
CC TIGR; HP0707; -.
CC HAMAP; MF_01007; -.
CC InterPro; IPR002903; Bac_Metnfrse.
CC Pfam; PF01795; Methyltransf_5; 1.
CC ProDom; PD004685; Bac_Metnfrse; 1.
CC TIGRFAMs; TIGR00006; TIGR00006; 1.
CC Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 308 AA; 34976 MW; 9C435C791E19FBB CRC64;

Query Match 1.8%; Score 6; DB 1; Length 308;
Best Local Similarity 100.0%; Pred.No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 EFLSSF 308
DB 186 EFLSSF 191

RESULT 207
DHL2_LACCO
ID DHL2_LACCO STANDARD; PRT; 309 AA.
AC P14295;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE L-2-hydroxyisocaproate dehydrogenase (EC 1.1.1.-) (L-HicDH).
OS Lactobacillus confusus.
OC Bacteria; Firmicutes; Lactobacillales; Weissella.
OX NCBI_TaxID=1583;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 20196;
RX MEDLINE=90060838; PubMed=2684788;
RA Lerch H.-P., Frank R., Collins J.;
RT "Cloning, sequencing and expression of the L-2-hydroxyisocaproate
RT dehydrogenase-encoding gene of Lactobacillus confusus in Escherichia
RT coli."
RL Gene 83:263-270(1989).
RN [2]
RP SEQUENCE OF 1-25.
RA Tsai H., Lerch H.-P., Kalwass H., Schuette H., Hoppe J., Collins J.;
RA (In) Neijssel O.M., van der Meer R.R., Luyben K.C.A.M. (eds.);
RL Proceedings 4th European congress biotechnology, pp.2:228-231,
RL Elsevier, Amsterdam (1987).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=95371125; PubMed=7643402;
RA Niefind K., Hecht H.-J., Schomburg D.;
RT "Crystal structure of L-2-hydroxyisocaproate dehydrogenase from
RT Lactobacillus confusus at 2.2-A resolution. An example of strong
RT asymmetry between subunits."
RL J. Mol. Biol. 251:256-281(1995).
CC -!- FUNCTION: Catalyzes the NADP dependent reversible and
CC stereospecific interconversion between 2-ketocarboxylic acids
CC and L-2-hydroxy-carboxylic acids. 2-ketoacids with medium chain
CC length (five to six C-atoms) are the best substrates.
CC -!- SUBUNIT: Homotetramer.
CC -!- MISCELLANEOUS: Can be applied in an industrial process for the
CC production of L-amino acid.
CC -!- SIMILARITY: Belongs to the LDH family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M31425; AAA88213.1; -.
CC PIR; JQ0114; JQ0114.
CC PDB; 1HYH; 15-OCT-95.
CC InterPro; IPR001557; L_LDH.
CC InterPro; IPR001236; 1dh.
CC InterPro; IPR000205; NAD_BS.
CC Pfam; PF00056; 1dh; 1.
CC Pfam; PF02866; 1dh_C; 1.
CC PRINTS; PR00086; LLDHGRNASE.
CC PROSITE; PS00064; L_LDH; 1.
CC Oxidoreductase; NADP; 3D-structure.
KW INIT MET 0 0 BY SIMILARITY.
FT ACT_SITE 178 178
FT STRAND 3 7
FT HELIX 11 23
FT TURN 24 24
FT STRAND 28 32
FT HELIX 36 49
FT HELIX 50 52
FT STRAND 58 61
FT HELIX 64 67

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

EMBL; AB065517; BAC05765.1; -  
Genew; HGNC:14743; OR4C6.  
InterPro; IPR000276; GPCR\_Rhodopsin.  
Pfam; PF00001; 7tm\_1; 1.  
PRINTS; PR00237; GPCRHOPOPSN.  
PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
G-protein coupled receptor; Transmembrane; Glycoprotein;  
Multigene family; Olfaction.  
DOMAIN 1 23 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 24 47 1 (POTENTIAL).  
FT DOMAIN 48 55 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 56 77 2 (POTENTIAL).  
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 99 118 3 (POTENTIAL).  
FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 138 156 4 (POTENTIAL).  
FT DOMAIN 157 193 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 194 217 5 (POTENTIAL).  
FT DOMAIN 218 233 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 234 256 6 (POTENTIAL).  
FT DOMAIN 257 267 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 268 287 7 (POTENTIAL).  
FT DOMAIN 288 309 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 289 309 BY SIMILARITY.  
FT CARBOHYD 6 N-LINKED (GLCNAC..) (POTENTIAL).  
SQ SEQUENCE 309 AA; 34557 MW; C481E2356F227426 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 309;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EXLQVS 219  
|||  
Db 289 EXLQVS 294

RESULT 208  
O4C6\_HUMAN STANDARD; PRT; 309 AA.  
ID O4C6\_HUMAN  
AC Q8NH72;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Olfactory receptor 4C6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,  
RA Tatemura S., Aburatani H., Asai K., Akiyama Y.;  
RT "Genome-wide discovery and analysis of human seven transmembrane helix  
receptor genes.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Putative odorant receptor.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC -!- DATABASE: NAME=Human Olfactory Receptor Data Explorer (HORDE);  
WWW="http://bioinformatics.weizmann.ac.il/cgi-bin/HORDE/showGene.pl?key=symbols";  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

EMBL; AB065517; BAC05765.1; -  
Genew; HGNC:14743; OR4C6.  
InterPro; IPR000276; GPCR\_Rhodopsin.  
Pfam; PF00001; 7tm\_1; 1.  
PRINTS; PR00237; GPCRHOPOPSN.  
PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
G-protein coupled receptor; Transmembrane; Glycoprotein;  
Multigene family; Olfaction.  
DOMAIN 1 23 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 24 47 1 (POTENTIAL).  
FT DOMAIN 48 55 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 56 77 2 (POTENTIAL).  
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 99 118 3 (POTENTIAL).  
FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 138 156 4 (POTENTIAL).  
FT DOMAIN 157 193 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 194 217 5 (POTENTIAL).  
FT DOMAIN 218 233 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 234 256 6 (POTENTIAL).  
FT DOMAIN 257 267 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 268 287 7 (POTENTIAL).  
FT DOMAIN 288 309 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 289 309 BY SIMILARITY.  
FT CARBOHYD 6 N-LINKED (GLCNAC..) (POTENTIAL).  
SQ SEQUENCE 309 AA; 34557 MW; C481E2356F227426 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 309;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
|||  
Db 194 GLLVTL 199

RESULT 209  
DHVS\_SULTO STANDARD; PRT; 311 AA.  
ID DHVS\_SULTO  
AC Q971T3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable deoxyhypusine synthase (EC 2.5.1.46) (DHS).  
GN DYS OR ST1293  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 10545 / 7;  
RX MEDLINE=21456156; PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Roki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermoacidophilic  
Crenarchaeon, Sulfolobus tokodaii strain 7.";  
RL DNA Res. 8:123-140(2001).  
CC -!- FUNCTION: Catalyzes the NAD-dependent oxidative cleavage of  
spermidine and the subsequent transfer of the butylamine moiety of  
spermidine to the epsilon-amino group of a specific lysine residue  
of the eIF-5A precursor protein to form the intermediate



```
RESULT 212
HEM3_PSESM
ID HEM3_PSESM STANDARD; PRT; 313 AA.
AC Q88B91;
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
DE synthase) (HMB) (Pre-uroporphyrinogen synthase).
GN HEMC OR PSPT00128.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22834015; PubMed=12928499;
RC STRAIN=DC3000;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -1- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
CC hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -1- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
CC hydroxymethylbilane + 4 NH(3).
CC -1- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
CC porphobilinogen subunits are added (By similarity).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the HMBs family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 277652; CAB01114.1; -
DR F1R; T18980; T18980.
DR WormPep; C06B3.4; CE07961.
DR InterPro; IPR002198; ADH short.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Steroid biosynthesis; Oxidoreductase; NADP;
KW Multigene family.
FT NP BIND 47 76 NADP (BY SIMILARITY).
FT ACT SITE 202 202 BY SIMILARITY.
FT SEQUENCE 314 AA; 34605 MW; C261F9ED72EAC245 CRC64;
SQ
Query Match 1.8%; Score 6; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 212 DYKLL 217
Db 114 DYKLL 119
RESULT 214
O5BH HUMAN
ID O5BH HUMAN STANDARD; PRT; 314 AA.
AC Q8NGF7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Olfactory receptor 5B17.
GN OSB17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Teutumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Putative odorant receptor.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- DATABASE: NAME=Human Olfactory Receptor Data Explorer (HORDE);
CC WWW="http://bioinformatics.weizmann.ac.il/cgi-bin/HORDE/showGene.pl?key=symbol
CC
Query Match 1.8%; Score 6; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 77 GLIVTL 82
Db 33 GLIVTL 38
RESULT 213
DHEV_CAEEL
ID DHEV_CAEEL STANDARD; PRT; 314 AA.
AC Q17703;
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB065849; BAC06067.1; --  
Genew; HGNC:15267; ORSB17.  
Pfam; PF00001; 7tm1.1;  
InterPro; IPR000276; GPCR\_Rhodopsin.  
PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
G-protein coupled receptor; Transmembrane; Glycoprotein;  
Multigene family; Olfaction.  
DOMAIN 1 23 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 24 44 1 (POTENTIAL).  
FT DOMAIN 45 52 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 53 73 2 (POTENTIAL).  
FT DOMAIN 74 97 3 (POTENTIAL).  
FT TRANSMEM 98 118 3 (POTENTIAL).  
FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 138 158 4 (POTENTIAL).  
FT DOMAIN 159 194 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 195 215 5 (POTENTIAL).  
FT DOMAIN 216 235 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 236 256 6 (POTENTIAL).  
FT DOMAIN 257 269 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 270 290 7 (POTENTIAL).  
FT DOMAIN 291 314 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 95 187 BY SIMILARITY.  
FT CARBOHYD 3 3 N-LINKED (GLCNAC... ) (POTENTIAL).  
SEQUENCE 314 AA; 35090 MW; 952366831PFD7052 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LVLTLI 83  
DB 209 LVLTLI 214

RESULT 215  
PRVA\_LISMO STANDARD; PRT; 314 AA.  
AC Q955A2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribosomal protein L11 methyltransferase (EC 2.1.1.-) (L11 Mtase).  
GN PRMA OR LM01471.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=10403S;  
RA Hanawa T., Kai M., Kaniya S., Yamamoto T.;  
RT "Cloning, sequencing, and transcriptional analysis of the dnaK heat shock operon of Listeria monocytogenes.";  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EGD-e / Serovar 1/2a;  
RX MEDLINE=21537279; PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Ruenick C., Amend A.,  
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Ettian K.-D., Fsihi H., Garcia-Del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordieck G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RT "Comparative genomics of Listeria species.";  
RL SCIENCE 294:849-852(2001).  
CC 1- FUNCTION: Methylates ribosomal protein L11 (By similarity).  
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC 1- SIMILARITY: Belongs to the methyltransferase superfamily. Prma family.  
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB023064; BAA82791.1; --  
EMBL; AL591979; CAC99549.1; --  
PIR; AG1258; AG1258.  
PIR; T43740; T43740.  
ListList; LMO01471; --  
HAMAP; MF\_00735; -- 1.  
InterPro; IPR004498; Ribosomal\_Prma.  
InterPro; IPR000051; SAM\_bind.  
TIGRFAMS; TIGR00406; prma; 1.  
KW Transferase; Methyltransferase; Complete proteome.  
SQ SEQUENCE 314 AA; 34811 MW; E71F4AF1DDF437F6 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSTFDI 179  
DB 88 LSTFDI 93

RESULT 216  
DHXX\_CAEEL STANDARD; PRT; 315 AA.  
AC O17795;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Putative steroid dehydrogenase F11A5.12 (EC 1.1.1.-).  
GN F11A5.12.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Gardner A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC 1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY. 17-BETA-HSD 3 SUBFAMILY.  
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z92830; CAB07363.1; --  
PIR; T20756; T20756.  
WormPep; F11A5.12; CE15790.

```

DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Steroid biosynthesis; Oxidoreductase; NADP;
KW Multigene family.
FT NP_BIND 47 76 NADP (BY SIMILARITY).
FT ACT_SITE 202 202 BY SIMILARITY.
SQ SEQUENCE 315 AA; 34872 MW; F4C9D3DBEFA539A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
DB 114 DYKLL 119

RESULT 217
MFTC_HUMAN STANDARD; PRT; 315 AA.
AC Q9H2D1; Q96JZ6; Q96S07;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mitochondrial folate transporter/carrier.
GN MFTC OR MFT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538421; PubMed=10978331;
RA Titus S.A., Moran R.G.;
RT "Retrovirally mediated complementation of the glyB phenotype. Cloning
RT of a human gene encoding the carrier for entry of folates into
RT mitochondria."
RL J. Biol. Chem. 275:36811-36817(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aoteuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saico K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

CC -!- FUNCTION: Transport folate across the inner membranes of mitochondria.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- SIMILARITY: Belongs to the mitochondrial carrier family.  
 CC -!- SIMILARITY: Contains 3 Solcar repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AF283645; AAG37834.1; --  
 CC EMBL; AK027531; BAB55180.1; --  
 CC EMBL; AK027787; BAB55368.1; --  
 CC EMBL; BC021893; AAH21893.1; --  
 CC GO; GO:0005743; C:mitochondrial inner membrane; NAS.  
 CC GO; GO:0008517; P:folate transporter activity; NAS.  
 CC GO; GO:0015884; P:folate transport; NAS.  
 CC InterPro: IPR001993; Mitoch carrier.  
 CC Pfam: PF00153; mito carr; 3.  
 CC PROSITE; PS50320; SOLCAR; 3.  
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.  
 FT TRANSMEM 89 106 POTENTIAL.  
 FT TRANSMEM 227 243 POTENTIAL.  
 FT TRANSMEM 281 300 POTENTIAL.  
 FT REPEAT 20 109 SOLCAR 1.  
 FT REPEAT 118 209 SOLCAR 2.  
 FT REPEAT 222 306 SOLCAR 3.  
 FT CONFLICT 117 117 R -> H (IN REF. 1).  
 FT CONFLICT 306 306 F -> L (IN REF. 2; BAB55368).  
 SQ SEQUENCE 315 AA; 35407 MW; EED376828B4D1069 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 TVEYIS 120  
 DB 224 TVEYIS 229  
 |||||  
 |||||

RESULT 218  
 MFTC MACFA  
 ID MFTC MACFA STANDARD; PRT; 315 AA.  
 AC Q95J75;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Mitochondrial folate transporter/carrier.  
 GN MFTC.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Transport folate across the inner membranes of  
 CC mitochondria (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -!- SIMILARITY: Belongs to the mitochondrial carrier family.

DR	EMBL; U39789; AAB17903.1; --.
DR	HSSP; P07862; LIOV.
DR	HMAP; MF 00047; --; 1.
DR	InterPro; IPR000291; Dala_lig_Van.
DR	Pfam; PF01820; Dala_Dala_ligas; 1.
DR	PROSITE; PS00843; DALA_DALA_LIGASE_1; 1.
DR	PROSITE; PS00844; DALA_DALA_LIGASE_2; 1.
KW	Ligase; Cell wall; Peptidoglycan synthesis.
FT	NON TER 316 316
SQ	SEQUENCE 316 AA; 35261 MW; 4FE5CDFDD716A8FB CRC64;
Query Match	
Best Local Similarity 100.0%; Pred.No.3.5e+02; Length 316;	
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	244 IMTKYI 249
Dd	126 IMTKYI 131
RESULT 220	
OTC_BACTN	
ID	OTC_BACTN STANDARD; PET; 318 AA.
AC	QBALES;
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Ornithine carbamoyltransferase (EC 2.1.1.3) (OTCase).
GN	ARGF OR BT317.
OS	Bacteroides thetaotaomicron.
OC	Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC	Bacteroidaceae; Bacteroides.
OX	NCBI_TaxID=818;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=VPI-5482 / ATCC 29148;
RX	MEDLINE=22550858; PubMed=12663928;
RA	Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA	Chiang H.C., Hooper L.V., Gordon J.I.;
RT	"A genomic view of the human-Bacteroides thetaotaomicron symbiosis.";
SC	Science 299:2074-2076(2003).
CC	-1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC	+ L-citrulline.
CC	-1- PATHWAY: Arginine biosynthesis; sixth step.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC	-1- SIMILARITY: Belongs to the ATCase/OTCase family.
CC	-1- CAUTION: Lacks the conserved threonine residue in position 48,
CC	which is part of the carbamoylphosphate binding site; it is
CC	replaced by a leucine residue.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AE016941; AAO78822.1; --.
DR	HMAP; MF 01109; atypical; 1.
DR	InterPro; IPR006130; Asp/Orn_Cotranf.
DR	InterPro; IPR006131; OTCase_O.
DR	InterPro; IPR006132; OTCase_P.
DR	Pfam; PF00185; OTCase; 1.
DR	Pfam; PF02729; OTCase_N; 1.
DR	PRINTS; PR00100; AOTCASE.
DR	PROSITE; PS00097; CARBAMOYLTRANSFERASE; FALSE NEG.
KW	Arginine biosynthesis; Transferase; Complete Proteome.
FT	SITE 47 51
FT	CARBAMOYLPHOSPHATE BINDING (BY
FT	SIMILARITY).
FT	SITE 110 110
FT	CARBAMOYLPHOSPHATE BINDING (BY
FT	SIMILARITY).
FT	SITE 147 147
FT	CARBAMOYLPHOSPHATE BINDING (BY



```
FT SITE 160 160 SIMILARITY).
FT FT IMPORTANT FOR STRUCTURAL INTEGRITY (BY
FT FT SIMILARITY).
FT SITE 273 276 ORNITHINE BINDING (BY SIMILARITY).
SQ SEQUENCE 318 AA; 36382 MW; 88116898BD42822B CRC64;

Query Match 1.8%; Score 6; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 FKVEL 174
Db 28 FKVEL 33

RESULT 221
XERC LEPIN
ID XERC LEPIN STANDARD; PRT; 318 AA.
AC Q7ZMB;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine recombinase xerc.
GN XERC OR LA2347.
OS Leptosira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptosira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
Xu J.-G., Zhao G.-P.;
RA "Unique physiological and pathogenic features of Leptosira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
CC -!- FUNCTION: Site-specific tyrosine recombinase, which acts by
catalyzing the cutting and rejoining of the recombining DNA
molecules. The xerc-xerD complex is essential to convert dimers of
the bacterial chromosome into monomers to permit their segregation
at cell division. It also contributes to the segregational
stability of plasmids (By similarity).
CC -!- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two
molecules of xerc and two molecules of xerD (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the "phage" integrase family. Xerc
subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AE011403; AAN49546.1;
CC HAMAP: MF 01808; 1;
CC InterPro: IPR004107; Phage_integr N.
CC Pfam: PF02899; Phage_integr N; 1.
CC Pfam: PF05889; Phage_integr; 1.
KW DNA recombination; DNA integration; Cell division;
Chromosome partition; DNA-binding; Complete proteome.
FT ACT_SITE 170 170 BY SIMILARITY.
FT ACT_SITE 194 194 BY SIMILARITY.
FT ACT_SITE 262 262 BY SIMILARITY.
FT ACT_SITE 265 265 BY SIMILARITY.
```

```
FT ACT_SITE 288 288 BY SIMILARITY.
FT ACT_SITE 297 297 TRANSIENT COVALENT LINKAGE TO DNA DURING
FT STRAND CLEAVAGE AND REJOINING (BY
FT SIMILARITY).
SQ SEQUENCE 318 AA; 37429 MW; 56F72FB4DE7C4C1B CRC64;

Query Match 1.8%; Score 6; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGL 78
Db 164 LYSSGL 169

RESULT 222
YQBD BACSU
ID YQBD BACSU STANDARD; PRT; 322 AA.
AC P45920;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yqbd.
GN YQBD OR BSU26150.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=95219086; PubMed=7704261;
RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
RT "Complete nucleotide sequence of a skin element excised by DNA
rearrangement during sporulation in Bacillus subtilis.";
RL Microbiology 141:323-327(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
the Bacillus subtilis genome containing the skin element and many
sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.Y., Glaser P., Goffeau A., Coliclighty E.J., Grandi G.,
Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Kluwe M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kasahara Y., Kluwe-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
Medina N., Meilado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Pario V., Pohl T.M., Portetle D., Portetle S., Prescott A.M.,
Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
Sorkin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpatra P., Tognoni A.,
```

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).  
RN [4]  
RP IDENTIFICATION.  
RX MEDLINE=96084975; PubMed=749895;  
RA Medigue C., Moszer I., Viari A., Danchin A.;  
RT "Analysis of a Bacillus subtilis genome fragment using a co-operative  
RT computer system prototype.";  
RL Gene 165:GC37-GC51(1995).  
CC -1- SIMILARITY: STRONG, TO B.SUBTILIS XKDF.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D32216; BAA06937.1; -  
DR EMBL; D84432; BAA12399.1; -  
DR EMBL; Z99117; CAB14556.1; -  
DR PIR; G69946; G69946.  
DR Subtilisin; BGI1275; yqbd.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 322 AA; 36247 MW; 9922F62EF6A000A CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 322;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 37 KASEEV 42  
Db 129 KASEEV 134  
|||||  
-----  
RESULT 223  
MC3R MOUSE  
ID MC3R MOUSE STANDARD; PRT; 323 AA.  
AC P33013;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Melanocortin-3 receptor (MC3-R).  
GN MC3R.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94226597; PubMed=8172596;  
RA Desarnaud F., Labbe O., Eggerickx D., Vassart G., Parmentier M.;  
RT "Molecular cloning, functional expression and pharmacological  
RT characterization of a mouse melanocortin receptor gene.";  
RL Biochem. J. 299:367-373(1994).  
CC -1- FUNCTION: Receptor for MSH (alpha, beta and gamma) and ACTH. This  
CC receptor is mediated by G proteins which activate adenylylate  
CC cyclase.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: Brain.  
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X74983; CAA52918.1; -  
DR PIR; S43850; S43850.  
DR MGD; MGI:96929; MC3R.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm.1.1.  
DR PRINTS; PRO0237; GPCRHOPOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
DR PROSITE; PS0262; G\_PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 38 63 1 (POTENTIAL).  
FT DOMAIN 64 75 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 76 100 2 (POTENTIAL).  
FT DOMAIN 101 118 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 119 140 3 (POTENTIAL).  
FT DOMAIN 141 160 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 161 181 4 (POTENTIAL).  
FT DOMAIN 182 186 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 187 210 5 (POTENTIAL).  
FT DOMAIN 211 245 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 246 268 6 (POTENTIAL).  
FT DOMAIN 269 277 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 278 301 7 (POTENTIAL).  
FT DOMAIN 302 323 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 2 2 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 16 16 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 28 28 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT LIPID 315 315 S-palmitoyl cysteine (POTENTIAL).  
SQ SEQUENCE 323 AA; 35806 MW; F4B7B02FA4A87B7B CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 50 KEILCG 55  
Db 311 KEILCG 316  
|||||  
-----  
RESULT 224  
MC3R RAT  
ID MC3R RAT STANDARD; PRT; 323 AA.  
AC P32244;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Melanocortin-3 receptor (MC3-R).  
GN MC3R.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Fischer; TISSUE=Hypothalamus;  
RX MEDLINE=9402273; PubMed=8415620;  
RA Roselli-Rehfuess L., Mountjoy K.G., Robbins L.S., Mortrud M.T.,  
RA Low M.J., Simerly R.B., Cone R.D.;  
RT "Identification of a receptor for gamma melanotropin and other  
RT proopiomelanocortin peptides in the hypothalamus and limbic system.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:8856-8860(1993).  
CC -1- FUNCTION: Receptor for MSH (alpha, beta and gamma) and ACTH.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: Brain.  
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X70667; CAAS0005.1; -;  
CC PIR; A48254; S36636;  
CC InterPro; IPR000276; GPCR\_Rhodospn.  
CC PRINTS; PR00237; GPCR\_RHODOPSIN.  
CC PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.  
CC PROSITE; PS0262; G PROTEIN RECEPTOR FL2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
CC Phosphorylation; Lipoprotein; Palmitate.  
CC DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 38 63 1 (POTENTIAL).  
CC DOMAIN 64 75 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 76 100 2 (POTENTIAL).  
CC DOMAIN 101 118 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 119 140 3 (POTENTIAL).  
CC DOMAIN 141 160 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 161 181 4 (POTENTIAL).  
CC DOMAIN 182 186 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 187 210 5 (POTENTIAL).  
CC DOMAIN 211 245 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 246 268 6 (POTENTIAL).  
CC DOMAIN 269 277 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 278 301 7 (POTENTIAL).  
CC DOMAIN 302 323 CYTOPLASMIC (POTENTIAL).  
CC CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC LIPID 315 315 S-palmitoyl cysteine (POTENTIAL).  
CC SEQUENCE 323 AA; 35866 MW; F4E9895C75E70A36 CRC64;  
-----  
Query Match 1.8%; Score 6; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 50 KEILCG 55  
DB 311 KEILCG 316  
-----  
RESULT 225  
RAF\_MSV36  
ID RAF\_MSV36 STANDARD; PRT; 323 AA.  
AC P00532;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serine/threonine-protein kinase transforming protein raf  
DE (EC 2.7.1.37).  
GN V-RAF.  
OS Murine sarcoma virus 3611.  
OC Viruses; Retroviridae; Retroviridae; Mammalian type C retroviruses.  
OX NCBI\_TaxID=11812;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84121298; PubMed=6320371;  
RA Kan N.C., Flordellis C.S., Mark G.E., Duesberg P.H., Papas T.S.;  
RT "A common onc gene sequence transduced by avian carcinoma virus MH2  
and by murine sarcoma virus 3611.";  
RL Science 223:813-816(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84172180; PubMed=6324342;  
RA Mark G.E., Rapp U.R.;  
RT "Primary structure of v-raf: relatedness to the src family of  
oncogenes.";  
RL Science 224:285-289(1984).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-RAF

CC POLYPROTEIN.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. RAF  
CC subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; K01691; AAA46579.1; ALT\_INIT.  
CC PIR; A00638; TMVUF6.  
CC HSP; P12931; 1FMK.  
CC InterPro; IPR00719; Prot\_kinase.  
CC InterPro; IPR008271; Ser\_thr\_kin\_AS.  
CC Pfam; PF00069; kinase; 1.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC KW Polyprotein; Serine/threonine-protein kinase; Transferase; Oncogene;  
CC KW ATP-binding.  
CC FT DOMAIN 24 284 PROTEIN KINASE  
CC NP\_BIND 30 38 ATP (BY SIMILARITY).  
CC BINDING 50 50 ATP (BY SIMILARITY).  
CC FT ACT SITE 143 143 BY SIMILARITY.  
CC SEQUENCE 323 AA; 36883 MW; 52A5423A66E362F3 CRC64;  
-----  
Query Match 1.8%; Score 6; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 22 KDNLA1 27  
DB 89 KDNLA1 94  
-----  
RESULT 226  
ATPT\_YEAST  
ID ATPT\_YEAST STANDARD; PRT; 325 AA.  
AC P22135;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ATP12 protein, mitochondrial precursor.  
GN ATP12 OR YUL180C OR J0486.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91210261; PubMed=1826907;  
RA Bowman S., Ackerman S.H., Griffiths D.E., Tzagoloff A.;  
RT "Characterization of ATP12, a yeast nuclear gene required for the  
assembly of the mitochondrial F1-ATPase.";  
RL J. Biol. Chem. 266:7517-7523(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Essential for the assembly of the mitochondrial  
CC F1-F0 complex.  
CC -!- SUBUNIT: Exists either as a homo- or heterooligomer.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL, EITHER AS A CONSTITUENT OF  
CC THE MATRIX, OR IN TENOUS ASSOCIATION WITH THE INTERNAL SIDE OF  
CC THE INNER MEMBRANE.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M61773; AAA34442.1; --  
DR EMBL; Z49455; CAA89475.1; --  
DR PIR; S56963; S56963.  
DR GerMOnline; 141792; --  
DR SGD; S0003716; ATP12.  
DR GO; GO:0003754; F:chaperone activity; IPI.  
DR GO; GO:0006461; P:protein complex assembly; IMP.  
KW Mitochondrion; Transic peptide.  
FT TRANSIT 1 732 MITOCHONDRION (POTENTIAL).  
FT CHAIN 733 325 ATP12 PROTEIN.  
FT CONFLICT 48 48 S -> N (IN REF. 1).  
SQ SEQUENCE 325 AA; 36554 MW; 0571C6C493E12CB1 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 25 LAILEX 30  
DB 242 LAILEX 247  
  
RESULT 227  
LACD LACLA STANDARD; PRT; 326 AA.  
AC P26533;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tagatose 1,6-diphosphate aldolase (EC 4.1.2.40) (Tagatose-bisphosphate  
DE aldolase) (D-tagatose-1,6-bisphosphate aldolase).  
GN LACD.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Plasmid pMG820.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91031107; PubMed=2125052;  
RA de Vos W.M., Boerrigter I., van Rooyen R.J., Reiche B.,  
RA Hengstenberg W.;  
RT "Characterization of the lactose-specific enzymes of the  
RT phosphotransferase system in Lactococcus lactis.";  
RL J. Biol. Chem. 265:22554-22560(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG1820;  
RX MEDLINE=91201377; PubMed=1901863;  
RA van Rooijen R.J., van Schalkwijk S., de Vos W.M.;  
RT "Molecular cloning, characterization, and nucleotide sequence of the  
RT tagatose 6-phosphate pathway gene cluster of the lactose operon of  
RT Lactococcus lactis.";  
RL J. Biol. Chem. 266:7176-7181(1991).  
CC -|- CATALYTIC ACTIVITY: D-tagatose 1,6-bisphosphate = glycerone  
CC phosphate + D-glyceraldehyde 3-phosphate.  
CC -|- PATHWAY: Tagatose 6-phosphate pathway of lactose catabolism.  
CC -|- SIMILARITY: Belongs to the aldolase lacD family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M61773; AAA34442.1; --  
DR EMBL; Z49455; CAA89475.1; --  
DR PIR; S56963; S56963.  
DR GerMOnline; 141792; --  
DR SGD; S0003716; ATP12.  
DR GO; GO:0003754; F:chaperone activity; IPI.  
DR GO; GO:0006461; P:protein complex assembly; IMP.  
KW Mitochondrion; Transic peptide.  
FT TRANSIT 1 732 MITOCHONDRION (POTENTIAL).  
FT CHAIN 733 325 ATP12 PROTEIN.  
FT CONFLICT 48 48 S -> N (IN REF. 1).  
SQ SEQUENCE 325 AA; 36554 MW; 0571C6C493E12CB1 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 25 LAILEX 30  
DB 242 LAILEX 247  
  
RESULT 227  
LACD LACLA STANDARD; PRT; 326 AA.  
AC P26533;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tagatose 1,6-diphosphate aldolase (EC 4.1.2.40) (Tagatose-bisphosphate  
DE aldolase) (D-tagatose-1,6-bisphosphate aldolase).  
GN LACD.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Plasmid pMG820.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91031107; PubMed=2125052;  
RA de Vos W.M., Boerrigter I., van Rooyen R.J., Reiche B.,  
RA Hengstenberg W.;  
RT "Characterization of the lactose-specific enzymes of the  
RT phosphotransferase system in Lactococcus lactis.";  
RL J. Biol. Chem. 265:22554-22560(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MG1820;  
RX MEDLINE=91201377; PubMed=1901863;  
RA van Rooijen R.J., van Schalkwijk S., de Vos W.M.;  
RT "Molecular cloning, characterization, and nucleotide sequence of the  
RT tagatose 6-phosphate pathway gene cluster of the lactose operon of  
RT Lactococcus lactis.";  
RL J. Biol. Chem. 266:7176-7181(1991).  
CC -|- CATALYTIC ACTIVITY: D-tagatose 1,6-bisphosphate = glycerone  
CC phosphate + D-glyceraldehyde 3-phosphate.  
CC -|- PATHWAY: Tagatose 6-phosphate pathway of lactose catabolism.  
CC -|- SIMILARITY: Belongs to the aldolase lacD family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M61773; AAA34442.1; --  
DR EMBL; Z49455; CAA89475.1; --  
DR PIR; S56963; S56963.  
DR GerMOnline; 141792; --  
DR SGD; S0003716; ATP12.  
DR GO; GO:0003754; F:chaperone activity; IPI.  
DR GO; GO:0006461; P:protein complex assembly; IMP.  
KW Mitochondrion; Transic peptide.  
FT TRANSIT 1 732 MITOCHONDRION (POTENTIAL).  
FT CHAIN 733 325 ATP12 PROTEIN.  
FT CONFLICT 48 48 S -> N (IN REF. 1).  
SQ SEQUENCE 325 AA; 36554 MW; 0571C6C493E12CB1 CRC64;

DR EMBL; M60447; AAA25180.1; --  
DR EMBL; M65190; AAA25171.1; --  
DR PIR; D39778; D39778.  
DR HAMAP; MF 00734; --; 1.  
DR InterPro; IPR007377; LacD.  
DR InterPro; IPR005927; LacD\_Gpos.  
DR Pfam; PF04274; LacD; 1.  
DR TIGRFAMs; TIGR01232; lacD; 1.  
KW Lactose metabolism; Lyase; Plasmid.  
SQ SEQUENCE 326 AA; 36476 MW; 37F22F556F47941B CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 326;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 195 KVLVAD 200  
DB 53 KVLVAD 58  
  
RESULT 228  
LACD STRPN STANDARD; PRT; 326 AA.  
ID LACD STRPN STANDARD; PRT; 326 AA.  
AC Q970L3;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tagatose 1,6-diphosphate aldolase (EC 4.1.2.40) (Tagatose-bisphosphate  
DE aldolase) (D-tagatose-1,6-bisphosphate aldolase).  
GN LACD OR SP1190.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC BAA-334 / TIGR4;  
MEDLINE=21357209; PubMed=11463916;  
RA Tetcelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
RA Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
RA Holt I.B., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
RT "Complete genome sequence of a virulent isolate of Streptococcus  
RT pneumoniae.";  
RL Science 293:498-506(2001).  
CC -|- CATALYTIC ACTIVITY: D-tagatose 1,6-bisphosphate = glycerone  
CC phosphate + D-glyceraldehyde 3-phosphate.  
CC -|- PATHWAY: Tagatose 6-phosphate pathway of lactose catabolism.  
CC -|- SIMILARITY: Belongs to the aldolase lacD family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AB007420; AAK75299.1; --  
DR PIR; B95138; B95138.  
DR TIGR; SP1130; --  
DR HAMAP; MF 00734; --; 1.  
DR InterPro; IPR007377; LacD.  
DR InterPro; IPR005927; LacD\_Gpos.  
DR Pfam; PF04274; LacD; 1.  
DR TIGRFAMs; TIGR01232; lacD; 1.  
KW Lactose metabolism; Lyase; Complete proteome.  
SQ SEQUENCE 326 AA; 36372 MW; 298CC7E188C37E0 CRC64;

```
Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 326;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 KVLVAD 200
DB 53 KVLVAD 58

RESULT 229
LACD_STR6 STANDARD; PRT; 326 AA.
AC Q8DP2;
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tagatose 1,6-diphosphate aldolase (EC 4.1.2.40) (Tagatose-bisphosphate
aldolase) [D-tagatose-1,6-bisphosphate aldolase].
GN LACD OR SPRI073.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2142945; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszczyk L.C., Burgett S.,
DeHoff B.S., Estrom S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
Gilmore R., Glass J.S., Kojima H., Kraft A.R., Lagace R.E.,
LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
McAhen S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Statrud P.L.,
Glass J.I.;
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RT J. Bacteriol. 183:5709-5717(2001).
CC -!- CATALYTIC ACTIVITY: D-tagatose 1,6-bisphosphate = glyceraldehyde
3-phosphate + D-glyceraldehyde 3-phosphate.
CC -!- PATHWAY: Tagatose 6-phosphate pathway of lactose catabolism.
CC -!- SIMILARITY: Belongs to the aldolase lacD family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; AB046464; BAB03343.1; -.
DR InterPro; IPR001963; VP7.
DR Pfam; PF00434; VP7; 1.
DR ProDom; PD000191; VP7; 1.
KW Coat protein; Transmembrane; Glycoprotein.
FT TRANSMEM 32 48
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 326 AA; 37226 MW; 77D999E6562E2687 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 326;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 PTEAVA 67
DB 86 PTEAVA 91

RESULT 231
Y067_CHLTR
ID Y067_CHLTR STANDARD; PRT; 326 AA.
AC Q9S529; O84070;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative metal-binding lipoprotein CT067 precursor.
DE Putative metal-binding lipoprotein CT067 precursor.
GN CT067.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=L2/434/Bu;
RX MEDLINE=99392470; PubMed=10463174;
RA Bannantine J.P., Rockey D.D.;
RT "Use of primate model system to identify Chlamydia trachomatis protein
antigens recognized uniquely in the context of infection.";
RL Microbiology 145:2077-2085(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
```

RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans;  
RL Chlamydia trachomatis";  
SC Science 282:754-759(1998).  
CC -!- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM  
CC CT067/CT068/CT069/CT070 FOR A METAL. METAL-BINDING COMPONENT.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
CC (Probable).  
CC -!- SIMILARITY: Belongs to the bacterial solute-binding protein family  
CC 9.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF077010; AAC35948.1; ALT\_INIT.  
CC EMBL; AE001281; AAC67658.1; -.  
CC PIR; E71561; E71561.  
CC PHCI-2DPAGE; Q9S529; -.  
CC InterPro; IPR006128; Lipoprotein 4.  
CC InterPro; IPR006127; SBP\_bac\_9.  
CC Pfam; PF01297; SBP\_bac\_9; 1.  
CC PRINTS; PR00690; ADHESN\_FAMILY.  
CC Hypothetical protein; Transposir; Metal-binding; Lipoprotein; Membrane;  
KW Signal; Complete proteome; Palmitate.  
FT SIGNAL 1 21  
FT CHAIN 22 326 PUTATIVE METAL-BINDING LIPOPROTEIN  
FT CT067.  
FT LIPID 22 22 N'-palmitoyl cysteine (Probable).  
FT LIPID 22 22 S'-diacylglycerol cysteine (Probable).  
FT VARIANT 190 190 A -> V (IN SEROVAR L2).  
SQ SEQUENCE 326 AA; 37035 MW; 5A5AA35AB6627D89 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 326;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 116 VEVYISA 121  
Db 254 VEVYISA 259  
RESULT 232  
ID YNZ9 CAEEL STANDARD; PRT; 326 AA.  
AC P45969;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein T09A5.9 in chromosome III.  
T09A5.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
PP SEQUENCE FROM N.A.  
RA STRAIN=Bristol N2;  
RA Thomas K.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: STRONG, TO S.POMBE AND YEAST SUS22.  
CC -!- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z36753; CAA85336.1; -.  
CC PIR; T24722; T24722.  
CC WormPep; T09A5.9; CE01090.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR007092; LRR SDS22.  
CC InterPro; IPR003603; LRRcap.  
CC Pfam; PF00560; LRR; 7.  
CC PRINTS; PR00019; LEURICHRPT.  
CC SMART; SM00446; LRRcap; 1.  
KW Hypothetical protein; Leucine-rich repeat; Repeat.  
FT REPEAT 35 57 LRR 1.  
FT REPEAT 58 80 LRR 2.  
FT REPEAT 81 102 LRR 3.  
FT REPEAT 103 126 LRR 4.  
FT REPEAT 128 146 LRR 5.  
FT REPEAT 147 170 LRR 6.  
FT REPEAT 172 190 LRR 7.  
FT REPEAT 191 212 LRR 8.  
FT REPEAT 213 236 LRR 9.  
FT REPEAT 238 256 LRR 10.  
FT REPEAT 257 280 LRR 11.  
FT REPEAT 281 304 LRR 12.  
SQ SEQUENCE 326 AA; 37359 MW; D4C5A5502FF0417B CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 326;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 174 LSTFDI 179  
Db 31 LSTFDI 36  
RESULT 233  
ALX STRPU  
ID ALX STRPU STANDARD; PRT; 327 AA.  
AC Q26657;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Aristalless homeobox protein (ALX) (SpPrx-1) (Fragment).  
GN ALX  
OS Strongylocentrotus purpuratus (Purple sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
OC Strongylocentrotus.  
OX NCBI\_TaxID=7668;  
RN [1]  
PP SEQUENCE FROM N.A.  
RA Martinez P., Davidson E.H.;  
RT "SpPrx-1, a sea urchin homeobox gene related to aristalless is  
RT expressed during embryogenesis";  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- SIMILARITY: Belongs to the paired homeobox family. Bicoid  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 homeobox domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D85080; BAA19774.1; -.  
CC HSRF; P06601; LFJL.  
CC InterPro; IPR001356; Homeobox.  
CC InterPro; IPR007104; Paired homeo.  
CC Pfam; PF00046; homeobox; 1.

```
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA_BIND 212 271 HOMEBOX.
FT DOMAIN 321 325 POLY-PRO.
FT NON_TER 327 327
SQ SEQUENCE 327 AA; 37146 MW; 0D387C5C72AECFD5 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 327;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 EAPQIA 139
Db 6 EAPQIA 11

RESULT 234
RPOA WIGBR
ID RPOA WIGBR STANDARD; PRT; 328 AA.
AC Q8DIY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (RNAP alpha
DE subunit) (Transcriptase alpha chain) (RNA polymerase alpha subunit).
GN RPOA OR WIGR5680.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia."
RL Nat. Genet. 32:402-407(2002).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: Homodimer. The RNAP catalytic core consists of 2 alpha, 1
CC beta, 1 beta', and 1 omega subunit (By similarity).
CC -!- DOMAIN: The N-terminal domain is essential for RNAP assembly and
CC basal transcription, whereas the C-terminal domain is involved in
CC interaction with transcriptional regulators and with upstream
CC promoter elements (By similarity).
CC -!- SIMILARITY: Belongs to the RNA polymerase alpha chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB063522; BAC24714.1; -.
CC HAMAP; MF 00059; -.
CC InterPro; IPR009025; RBP11-like_RNapo.
CC InterPro; IPR001700; RNA_polA_bac_org.
CC Pfam; PF01000; RNA_pol_A_bac; 1.
CC Pfam; PF03118; RNA_pol_A_CTD; 1.
CC ProDom; PD001179; RNA_polA_bac_org; 1.
CC SMART; SM00662; RPOLD1.
CC Complete proteome.

DR HAMAP; MF 00059; -.
DR InterPro; IPR009025; RBP11-like_RNapo.
DR InterPro; IPR001700; RNA_polA_bac_org.
DR Pfam; PF01000; RNA_pol_A_bac; 1.
DR Pfam; PF03118; RNA_pol_A_CTD; 1.
DR ProDom; PD001179; RNA_polA_bac_org; 1.
DR SMART; SM00662; RPOLD1.
KW Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
```

```
FT DOMAIN 1 233 ALPHA N-TERMINAL DOMAIN (ALPHA-NTD) (BY
FT SIMILARITY).
FT DOMAIN 247 328 ALPHA C-TERMINAL DOMAIN (ALPHA-CTD) (BY
FT SIMILARITY).
SQ SEQUENCE 328 AA; 36927 MW; 9876CE4C9AF98C67 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295
Db 78 IVEILL 83

RESULT 235
GLK_BACHD
ID GLK_BACHD STANDARD; PRT; 330 AA.
AC Q9KCZ4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucokinase (EC 2.7.1.2) (Glucose kinase).
GN GLK OR GLK OR BHI425.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- CATALYTIC ACTIVITY: ATP + D-glucose = ADP + D-glucose 6-phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ROK (NAGC/XYL) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF001512; BAB05144.1; -.
CC PIR; A83828; A83828.
CC InterPro; IPR000600; ROK.
CC InterPro; IPR004654; ROK_glcA_fam.
CC Pfam; PF00480; ROK; 1.
CC TIGRFAWS; TIGR00744; ROK_glcA_fam; 1.
CC PROSITE; PS01125; ROK; 1.
CC Transferase; Kinase; Glycolysis; ATP-binding; Complete proteome.
SQ SEQUENCE 330 AA; 34496 MW; E5448D6C8D87456F CRC64;

Query Match 1.8%; Score 6; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 ASDAFA 185
Db 231 ASDAFA 236

RESULT 236
G3P2_KLUMA
ID G3P2_KLUMA STANDARD; PRT; 331 AA.
AC Q01077;
DT 01-NOV-1997 (Rel. 35, Created)
```



DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase 2 (EC 1.2.1.12) (GAPDH 2).  
GN GAP2.  
OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
OX NCBI\_TaxID=4911;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 10022;  
RA MEDLINE=95397590; PubMed=7668042;  
RX Fernandes P.A., Sena-Estevés M., Moradas-Ferreira P.;  
RT "Characterization of the glyceraldehyde-3-phosphate dehydrogenase  
RT gene family from Kluyveromyces marxianus -- polymerase chain  
RT reaction-single-strand conformation polymorphism as a tool for the  
RT study of multigenic families.";  
RL Yeast 11:725-733(1995).  
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.  
CC -1- PATHWAY: Second phase of glycolysis; first step.  
CC -1- SUBUNIT: Homotetramer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate  
CC dehydrogenase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: S80241; AAC35203.1; -;  
CC PIR: S57280; S57280.  
CC HSSP: P06977; 1GAD.  
CC InterPro: IPR000173; GAP dhdrogenase.  
CC InterPro: IPR006424; GAPDH-I.  
CC Pfam: PF00044; gpdh; 1.  
CC Pfam: PF02800; gpdh; C; 1.  
CC PRINTS: PR00078; G3PFDHGNASE.  
CC TIGRFAMS: TIGR01534; GAPDH; 1.  
CC PROSITE: PS00071; GAPDH; 1.  
CC GlycoVisis; 149 149 GLYCERALDEHYDE 3-PHOSPHATE.  
FT BINDING 149 149 ACTIVATES THIOL GROUP DURING CATALYSIS.  
FT ACT\_SITE 176 176  
SQ SEQUENCE 331 AA; 35539 MW; FEDB08475F4B5F09 CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 154 PLAKII 159  
Db 156 PLAKII 161  
RESULT 237  
LDHD\_TREPA  
ID LDHD\_TREPA STANDARD; PRT; 331 AA.  
AC O83080;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE D-lactate dehydrogenase (EC 1.1.1.26) (D-LDH) (D-specific D-2-  
DE hydroxyacid dehydrogenase).  
GN LDHD OR TP0037.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=Nichols;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren B., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
RT spirochete.";  
RL Science 281:375-388(1998).  
CC -1- CATALYTIC ACTIVITY: (R)-lactate + NAD(+) = pyruvate + NADH.  
CC -1- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid  
CC dehydrogenase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AB001189; AAC65033.1; -;  
CC PIR: D71373; D71373.  
CC HSSP: P30901; 2DLQ.  
CC TIGR: TP0037; -;  
CC InterPro: IPR006139; 2-Hacid DH.  
CC InterPro: IPR006140; 2-Hacid DH\_C.  
CC InterPro: IPR000205; NAD BS.  
CC Pfam: PF00389; 2-Hacid DH; 1.  
CC Pfam: PF02826; 2-Hacid DH\_C; 1.  
CC PROSITE: PS00065; D\_2-HYDROXYACID\_DH\_1; 1.  
CC PROSITE: PS00670; D\_2-HYDROXYACID\_DH\_2; 1.  
CC PROSITE: PS00671; D\_2-HYDROXYACID\_DH\_3; 1.  
CC Oxidoreductase; NAD; Complete proteome.  
FT NP\_BIND 148 176 NAD (BY SIMILARITY).  
FT ACT\_SITE 235 235 SUBSTRATE BINDING (BY SIMILARITY).  
FT ACT\_SITE 264 264 BY SIMILARITY.  
FT ACT\_SITE 296 296 BY SIMILARITY.  
SQ SEQUENCE 331 AA; 36873 MW; 773B01E6E2384E0A CRC64;  
Query Match 1.8%; Score 6; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 129 LLKGYE 134  
Db 42 LLKGYE 47  
RESULT 238  
SYW\_UREPA  
ID SYW\_UREPA STANDARD; PRT; 333 AA.  
AC Q9PQW8;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TrpRS).  
GN TRPS OR UUI175.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serovar 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
RT urealyticum.";

RL Nature 407:757-762 (2000).

CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

CC diphosphate + L-tryptophanyl-tRNA(Trp).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF002117; AAF30582.1; -.

DR HSSP; P00953; 1D2R.

DR HAMAP; MF\_00140; -; 1.

DR InterPro; IPR002305; tRNA-synt\_1b.

DR InterPro; IPR001412; tRNA-synt\_1.

DR InterPro; IPR002306; Trp tRNA-synt\_1b.

DR Pfam; PF00579; tRNA-synt\_1b; 1.

DR PRINTS; PR01039; TRNASYNTHTRP.

DR TIGRFAMs; TIGR00233; trps; 1.

DR PROSITE; PS00178; AA TRNA LIGASE I; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW Complete proteome.

FT SITE 10 18 "HIGH" REGION

FT SITE 200 204 "RMSKS" REGION

FT BINDING 203 203 ATP (BY SIMILARITY).

SQ SEQUENCE 333 AA; 37755 MW; AE6646EB161797FF CRC64;

Query Match 1.8%; Score 6; DB 1; Length 333;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 VADFLE 203

DB 284 VADFLE 289

|||||

RESULT 239

FCN1\_MOUSE

ID FCN1\_MOUSE STANDARD; PRT; 334 AA.

AC 070165;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Ficolin 1 precursor (Collagen/fibrinogen domain-containing protein 1)

DE (Ficolin-A) (Ficolin A) (M-Ficolin).

GN FCN1 OR FCNA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Liver;

RX MEDLINE=98205801; PubMed=9535745;

RA Fujimori Y., Harumiya S., Fukumoto Y., Miura Y., Yagasaki K.,

RA Tachikawa H., Fujimoto D.;

RT "Molecular cloning and characterization of mouse ficolin-A";

RL Biochem. Biophys. Res. Commun. 244:796-800(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,

RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Trinchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: Involved in serum exerting lectin activity. Binds GlcNAc

CC (By similarity).

CC -!- SUBUNIT: HOMOPOLYMER. INTERACTS WITH ELASTIN (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: SECRETED. FOUND ON THE MONOCYTE SURFACE (BY

CC SIMILARITY).

CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER AND SPLEEN.

CC -!- SIMILARITY: Belongs to the ficolin lectin family.

CC -!- SIMILARITY: Contains 1 collagenous domain.

CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AB007813; BAA25126.1; -.

DR EMBL; BC019180; AAH19180.1; -.

DR PIR; JCS980; JCS980.

DR HSSP; P02671; 1FZD.

DR MGD; MGI:1340905; Fcna.

DR InterPro; IPR008161; Clg\_helix.

DR InterPro; IPR008160; Collagen.

DR InterPro; IPR002181; Fibrinogen\_C.

DR Pfam; PF01391; Collagen; 1.

DR Pfam; PF00147; fibrinogen\_C; 1.

DR ProDom; PD000007; Clg\_helix; 1.

DR SMART; SM00186; FBG; 1.

DR PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.

DR Lectin; Collagen; Repeat; Glycoprotein; Signal; Multigene family.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 334 FICOLIN 1.

FT DOMAIN 50 88 COLLAGEN-LIKE.

FT DOMAIN 152 298 FIBRINOGEN C-TERMINAL.

FT CARBOHYD 261 261 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 334 AA; 36298 MW; 9D30C05036AA04B1 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 334;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 KDLLTR 193

DB 127 KDLLTR 132

|||||

RESULT 240

G3P\_PICCI

ID G3P\_PICCI STANDARD; PRT; 334 AA.

AC Q9UVC0;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).

GN GPD.

OS Pichia ciferrii (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.

```

OX NCBI_TaxID=36020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14091;
RA Bae J.-H., Sohn J.-H., Choi E.-S., Park J.-S., Rhee S.-K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF053300; AAF21710.1; -.
CC HSP; P00357; 4GPD.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC Pfam; PF006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC PRINTS; PR00780; G3PDHGRGNASE.
CC TIGRfams; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD.
CC BINDING 151 151
CC ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
CC SEQUENCE 334 AA; 36180 MW; 6C29C3BD06C79599 CRC64;
DR EMBL; AF053300; AAF21710.1; -.
DR HSP; P00357; 4GPD.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00780; G3PDHGRGNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Glycolysis; Oxidoreductase; NAD.
DR BINDING 151 151
DR ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
DR SEQUENCE 334 AA; 36180 MW; 6C29C3BD06C79599 CRC64;
SQ
Query Match 1.8%; Score 6; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 154 PLAKII 159
DB 158 PLAKII 163
RESULT 241
ID ILVC_METJA STANDARD; PRT; 334 AA.
AC Q58938;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ketol-acid reductoisomerase [EC 1.1.1.86] (Acetohydroxy-acid
DE isomeroreductase) (Alpha-keto-beta-hydroxyacyl reductoisomerase).
GN ILVC OR MJ1543.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii.";
```

```

RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.
CC -1- PATHWAY: Valine and isoleucine biosynthesis; second step.
CC -1- SIMILARITY: Belongs to the ketol-acid reductoisomerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67595; AAB99561.1; ALT_INIT.
CC TIGR; M01543; -.
CC HAMAP; MF_00435; -. 1.
CC InterPro; IPR008927; 6GDGH_C like.
CC InterPro; IPR000506; Ach_isomrdctse.
CC Pfam; PF01450; ILVC; 1.
CC TIGRfams; TIGR00465; ilvc; 1.
CC Oxidoreductase; Branched-chain amino acid biosynthesis; NADP;
CC Complete proteome.
CC ACT SITE 112 112 POTENTIAL.
CC SEQUENCE 334 AA; 37300 MW; 9EC8DE89143B4002 CRC64;
SQ
Query Match 1.8%; Score 6; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 48 AMKEIL 53
DB 278 AMKEIL 283
RESULT 242
ID DBX1_MOUSE STANDARD; PRT; 335 AA.
AC P52950;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein DBX1.
GN DBX1 OR DBX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95110746; PubMed=7811640;
RA Lu S., Wise T.L., Ruddle F.H.;
RT "Mouse homeobox gene Dbx: sequence, gene structure and expression
RT pattern during mid-gestation."
RL Mech. Dev. 47:187-195(1994).
CC -1- FUNCTION: Could have a role in patterning the central nervous
CC system during embryogenesis.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: During early and mid-gestation, dbx
CC expression is restricted to the telencephalon, diencephalon,
CC dorsal mesencephalon and spinal cord. At later gestational stages,
CC dbx expression continues in the dorsal mesencephalon and
CC diencephalon, in which expression is more restricted than at the
CC earlier stages.
CC -----
CC -1- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

-----  
CC EMBL; S75837; AAB33013.1; --  
DR HSP; P23441; IFTT.  
DR MGD; MGI:94867; Dlx1.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR000047; HTH laubrepressr.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRINTS; PR00031; HTHREPRESSR.  
DR ProDom; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR InterPro; IPR00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.  
FT DOMAIN 181 240 HOMEBOX.  
FT BIND 319 331 ASP/GLU-RICH (HIGHLY ACIDIC).  
SQ SEQUENCE 335 AA; 36333 MW; 06312CD41F03F369 CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 247 KYISK 252  
Db 204 KYISK 209  
-----  
RESULT 243  
FCN1\_RAT  
ID FCN1\_RAT STANDARD; PRT; 335 AA.  
AC Q9WTS8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ficolin 1 precursor (Collagen/fibrinogen domain-containing protein 1)  
DE (Ficolin-A) (Ficolin A) (M-Ficolin).  
GN FCN1 OR FCNA.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RA Yoshida Y., Tachikawa H., Fujimori Y., Miura Y., Yagasaki K.,  
RA Fujimoto D., Harumiya S.;  
RT Molecular cloning and characterization of rat ficolin-A.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Involved in serum exerting lectin activity. Binds GlcNAc  
CC (BY similarity).  
CC -!- SUBUNIT: HOMOPOLYMER. INTERACTS WITH ELASTIN (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED. FOUND ON THE MONOCYTE SURFACE (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: Belongs to the ficolin lectin family.  
CC -!- SIMILARITY: Contains 1 collagenous domain.  
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AB026057; BAA76940.2; --  
DR HSP; P02671; IFZD.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR Pfam; PF01391; Collagen; 1.  
DR Pfam; PF00147; fibrinogen\_C; 1.  
DR SMART; SM00186; FBG; 1.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.

-----  
KW Lectin; Collagen; Repeat; Glycoprotein; Signal; Multigene family.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 335 FICOLIN 1.  
FT DOMAIN 50 88 COLLAGEN-LIKE.  
FT DOMAIN 152 298 FIBRINOGEN C-TERMINAL.  
FT CARBOHYD 271 271 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 335 AA; 36627 MW; 1A7FC9568E76ED5D CRC64;  
  
Query Match 1.8%; Score 6; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 188 KDLTR 193  
Db 127 KDLTR 132  
-----  
RESULT 244  
KIME\_PVRAB  
ID KIME\_PVRAB STANDARD; PRT; 335 AA.  
AC Q9V187;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Mevalonate kinase (EC 2.7.1.36) (MK).  
GN MVK OR PYRAB05410 OR PAB0372.  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GES / Orsay;  
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
RA Poch O., Priet D., Querellou J., Ripp R., Thierry J.-C.,  
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;  
RT "An integrated analysis of the genome of the hyperthermophilic  
archaeon Pyrococcus abyssi";  
RL Mol. Microbiol. 47:1495-1512(2003).  
CC -!- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-  
CC phosphate.  
CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the GMP kinase family. Mevalonate kinase  
CC subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AJ248284; CAB49463.1; --  
DR PIR; H75172; H75172.  
DR HAMAP; MF\_00217; --; 1.  
DR InterPro; IPR001174; Galkinase.  
DR InterPro; IPR006204; GMP kinase.  
DR InterPro; IPR006203; GMPKase ATP.  
DR InterPro; IPR006205; Mv\_gal\_kin.  
DR InterPro; IPR006206; Mv\_galkinase.  
DR Pfam; PF00288; GMP\_kinases; 1.  
DR PRINTS; PR00960; LMBPPROTEIN.  
DR PRINTS; PR00959; MEVGALKINASE.  
DR TIGRFAMs; TIGR00549; mevalon kin; 1.  
DR PROSITE; PS00627; GMP\_KINASES\_ATP; 1.  
KW Transferase; Kinase; ATP-binding; Magnesium; Complete proteome.  
FT BIND 111 121 ATP (POTENTIAL).  
SQ SEQUENCE 335 AA; 35774 MW; ED0B06EDA186599C CRC64;

Query Match 1.8%; Score 6; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKILF 161  
DB 11 AKILF 16

RESULT 245  
TALI HUMAN STANDARD; PRT; 337 AA.  
AC P27837; 000751;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Transaldolase [EC 2.2.1.2]  
GN TALDOI OR TALDO OR TAL OR TALDOR.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94132057; PubMed=8300619;  
RA Banki K., Halladay D., Perl A.;  
RT "Cloning and expression of the human gene for transaldolase. A novel  
RT highly repetitive element constitutes an integral part of the coding  
RT sequence.";  
RL J. Biol. Chem. 269:2847-2851(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97480738; PubMed=9339383;  
RA Banki K., Eddy R.L., Shows T.B., Halladay D.L., Bullrich F.,  
RA Croce C.M., Jurecic V., Baldini A., Perl A.;  
RT "The human transaldolase gene (TALDOI) is located on chromosome 11 at  
RT p15.4-p15.5";  
RL Genomics 45:233-238(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98192510; PubMed=9524206;  
RA Kusuda J., Hirai M., Toyoda A., Tanuma R., Nomura-Kitabayashi A.,  
RA Hashimoto K.;  
RT "Cloning and chromosomal localization of a paralog and a mouse homolog  
RT of the human transaldolase gene.";  
RL Gene 209:13-21(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20167206; PubMed=10702296;  
RA Perl A., Colombo E., Samoilova E., Butler M.C., Banki K.;  
RT "Human transaldolase-associated repetitive elements are transcribed by  
RT RNA polymerase III.";  
RL J. Biol. Chem. 275:7261-7272(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP VARIANT TALDOI DEFICIENCY SER-171 DEL.  
RX MEDLINE=21205587; PubMed=11283793;  
RA Verhoeven N.M., Huck J.H.J., Roos B., Struys E.A., Salomons G.S.,  
RA Douwes A.C., van der Knaap M.S., Jakobs C.;  
RT "Transaldolase deficiency: liver cirrhosis associated with a new  
RT inborn error in the pentose phosphate pathway.";  
RL Am. J. Hum. Genet. 68:1086-1092(2001).  
CC -I- FUNCTION: Transaldolase is important for the balance of  
CC metabolites in the pentose-phosphate pathway.  
CC -I- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde  
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.  
CC -I- PATHWAY: Pentose phosphate pathway; nonoxidative part.  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -I- DISEASE: Defects in TALDOI are a cause of a deficiency that  
CC results in telangiectases of the skin, hepatosplenomegaly, and  
CC enlarged clitoris.  
CC -I- SIMILARITY: Belongs to the transaldolase family. Subfamily 1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; LJ9437; AAB53943.1; -;  
DR EMBL; AF010400; AAC52068.1; -;  
DR EMBL; AF010398; AAC52068.1; JOINED.  
DR EMBL; AF010399; AAC52068.1; JOINED.  
DR EMBL; AF058913; AAF04078.1; -;  
DR EMBL; BC010103; AAH10103.1; -;  
DR PIR; A49985; A49985.  
DR PDB; 1F05; 13-JUL-00.  
DR Genew; HGNC:11559; TALDOI.  
DR GK; P37837; -;  
DR MIM; 602063; -;  
DR MIM; 606003; -;  
DR GO; GO:0004801; P:transaldolase activity; TAS.  
DR GO; GO:0005975; P:carbohydrate metabolism; TAS.  
DR InterPro; IPR001585; Transaldolase.  
DR InterPro; IPR004730; Transaldolase\_AB.  
DR Pfam; PF00923; Transaldolase; 1.  
DR TIGRFAMs; TIGR00874; talAB; 1.  
DR PROSITE; PS00958; TRANSALDOLASE 2; 1.  
DR PROSITE; PS01054; TRANSALDOLASE 1; 1.  
KW Transferase; Pentose shunt; Disease mutation; 3D-structure.  
FT ACT\_SITE 142 142 BY SIMILARITY  
FT VARIANT 171 171 MISSING (IN TALDOI DEFICIENCY).  
FT /FTID=VAR\_011511.  
FT CONFLICT 209 213 LEDPG -> WKTG (IN REF. 1).  
FT SEQUENCE 337 AA; 37540 MW; 8CB4992AEF364B64 CRC64;  
SQ  
Query Match 1.8%; Score 6; DB 1; Length 337;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLIGEL 235  
DB 258 KLIGEL 263

RESULT 246  
TALI MOUSE STANDARD; PRT; 337 AA.  
AC Q93052; P70358; P70703;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)





```
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RL genome between the gnt and iol operons.";
RN DNA Res. 2:61-69(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broutillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,
RA Hilbert H., Hoisappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Tepstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis.";
RN Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005554; BAA21586.1; -
DR EMBL; Z99124; CAB16035.1; -
DR PIR; F70071; F70071.
DR SubtilList; BG11109; yxaG.
DR InterPro; IPR007113; Cupin sup.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 337 AA; 37584 MW; F9F3255C98C215A CRC64;

Query Match 1.8%; Score 6; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 ASEEVS 43
DB 143 ASEEVS 148
|||||
RESULT 249
HRCAL STRAL STANDARD; PRT; 338 AA.
ID HRCAL STRAL STANDARD; PRT; 338 AA.

SEQUENCE FROM N.A.
STRAIN=A3(2) / M145;
MEDLINE=21996410; PubMed=12000953;
Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieffer H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieffer T., Larke L., Murphy J., Oliver K., O'Neil S.,
Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
```

```
AC OS2163;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat-inducible transcription repressor hrca.
GN HRCAL.
OS Streptomyces albus G.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1962;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J1074;
RX MEDLINE=98422466; PubMed=9748446;
RA Grandvalet C., Rapoport G., Mazodier P.;
RT "hrca, encoding the repressor of the groEL genes in Streptomyces
RT albus G, is associated with a second dnaJ gene.";
RL Bacteriol. 180:5129-5134(1998).
CC -I- FUNCTION: Negative regulator of class I heat shock genes (grpE-
CC dnaK-dnaJ and groELS operons). Prevents heat-shock induction of
CC these operons (By similarity).
CC -I- SIMILARITY: Belongs to the hrca family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF025656; AAC62528.1; -
DR HAMAP; MF 00081; -; 1.
DR InterPro; IPR002571; HrcA.
DR Pfam; PF01628; HrcA; 1.
DR TIGRFAMs; TIGR00331; hrca; 1.
KW Transcription regulation; Repressor; Heat shock.
SQ SEQUENCE 338 AA; 36976 MW; 81C15F9A92BA7A49 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 338;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKJLGE 234
DB 263 LKJLGE 268
|||||
RESULT 250
HRCAL STRCO STANDARD; PRT; 338 AA.
ID HRCAL STRCO STANDARD; PRT; 338 AA.
AC Q9RDD6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heat-inducible transcription repressor hrca.
GN HRCAL OR SCO2555 OR SCC77.22C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieffer H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieffer T., Larke L., Murphy J., Oliver K., O'Neil S.,
Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
```



```

RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: Negative regulator of class I heat shock genes (grpE-
CC dnaK-dnaJ and groELS operons). Prevents heat-shock induction of
CC these operons (By similarity).
CC -!- SIMILARITY: Belongs to the hrca family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL939113; CAB66233.1; -.
DR HAMAP; MF_00081; -.
DR InterPro; IPR002571; Hrca.
DR Pfam; PF01628; Hrca; 1.
DR TIGRFAMs; TIGR00331; hrca; 1.
KW Transcription regulation; Repressor; Heat shock; Complete proteome.
SQ SEQUENCE 338 AA; 36575 MW; 12BD83F913E9F871 CRC64;

Query Match      1.8%; Score 6; DB 1; Length 338;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLLGE 234
Db 263 LKLLGE 268

```

Search completed: April 12, 2004, 10:35:48  
 Job time : 29 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:32:28 ; Search time 46 Seconds  
(without alignment)  
2311.514 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 337

Sequence: 1 MKKWLFSKSHKNPAEIVKI.....FADEKNVLIKQIRDLKKTAP 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

SPTREMBL.25.\*  
1: sp archaea.\*  
2: sp bacteria.\*  
3: sp fungi.\*  
4: sp human.\*  
5: sp invertebrate.\*  
6: sp mammal.\*  
7: sp mhc.\*  
8: sp organelle.\*  
9: sp phage.\*  
10: sp plant.\*  
11: sp rodent.\*  
12: sp virus.\*  
13: sp vertebrate.\*  
14: sp unclassified.\*  
15: sp virus.\*  
16: sp bacteriaph.\*  
17: sp archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	244	72.4	289	4	Q96FG1
2	54	16.0	103	11	Q8K038
3	23	6.8	205	11	Q8K312
4	23	6.8	341	4	Q724X0
5	23	6.8	341	11	Q8VD28
6	17	5.0	343	13	Q803V8
7	14	4.2	377	5	Q722A5
8	14	4.2	636	5	Q21643
9	9	2.7	370	3	Q873K5
10	8	2.4	251	16	Q8DPL9
11	8	2.4	339	5	Q8M226
12	8	2.4	442	5	Q95333
13	8	2.4	464	5	Q9W107
14	8	2.4	576	16	Q87TU1
15	8	2.4	695	3	Q9P8T6
16	8	2.4	747	5	O76750

Q9V4R2 drosophila  
Q8R14 encephalito  
Q8VH6 mus musculu  
Q8MZQ7 drosophila  
Q94K21 arabidopsis  
Q81U84 bacillus an  
Q81H15 bacillus ce  
Q88AQ4 pseudomonas  
Q812Y6 sulfolobus  
Q8931 mycobacteri  
Q81K2 caenorhabdi  
Q82147 arabidopsis  
Q86F7 trypanosoma  
Q89Y02 staphylococ  
Q89X11 clostridium  
Q89W7 staphylococ  
Q82Q6 streptomyce  
Q8YK8 anabaena sp  
Q8YK2 anabaena sp  
Q8309 arabidopsis  
Q88W2 lactobacill  
Q8VX3 drosophila  
Q8AJD5 pseudomonas  
Q8XJG8 linaria vul  
Q8XJG8 linaria mar  
Q82C12 listeria in  
Q80I1 theiler-lik  
Q8XU5 oryza sativ  
Q8AB4 bacteroides  
Q8K28 aminomona  
Q83187 staphylococ  
Q82L17 rhizobium m  
Q86635 aquifex aeo  
Q89U1 bradyrhizob  
Q8EM80 oceanobacil  
Q85132 mus musculu  
Q87175 vibrio para  
Q8N7X8 homo sapien  
Q86108 homo sapien  
Q8AD8 homo sapien  
Q8816 rhizobium 1  
Q86MV4 homo sapien  
Q8XQ7 oryza sativ  
Q81347 dirofilaria  
Q86614 mus musculu  
Q88J4 shigella fl  
Q8P19 streptomyce  
Q8BVX5 homo sapien  
Q8X20 mus musculu  
Q8VX8 plutella xy  
Q87H6 vibrio para  
Q87E1 mus musculu  
Q8F91 corynebacte  
Q8BVX5 mus musculu  
Q81798 caenorhabdi  
Q8U4F8 synecococ  
Q84B8 linaria can  
Q84B7 linaria vul  
Q8BK7 mus musculu  
Q8X793 mycobacteri  
Q53442 mycobacteri  
Q7U0P0 mycobacteri  
Q52630 clostridium  
Q81Y16 homo sapien  
Q8B186 mus musculu  
Q87M91 clostridium  
Q8DPF4 streptococ  
Q89U7 staphylococ  
Q74388 schizosacch  
Q89402 paramacium  
Q8X5K0 escherichia  
Q7ZA52 magnaporthe

90	7	2.1	314	2	O87260	O87260 lactococcus	163	7	2.1	459	3	O875W3	O875W3 saccharomyc
91	7	2.1	314	2	O32778	O32778 lactococcus	164	7	2.1	475	16	O8XL24	O8XL24 clostridium
92	7	2.1	315	17	O970H0	O970H0 sulfolobus	165	7	2.1	476	10	O9AWX3	O9AWX3 oryza sativ
93	7	2.1	320	16	O06996	O06996 bacillus su	166	7	2.1	480	16	Q7VQI7	Q7VQI7 candidatus
94	7	2.1	321	3	O84BB3	O84BB3 erwinia chr	167	7	2.1	485	4	Q9H677	Q9H677 homo sapien
95	7	2.1	322	3	O74237	O74237 candida ten	168	7	2.1	489	11	Q924K5	Q924K5 rattus norv
96	7	2.1	322	10	O8LIF3	O8LIF3 oryza sativ	169	7	2.1	490	6	Q726Q9	Q726Q9 homo sapien
97	7	2.1	332	17	O97YG7	O97YG7 sulfolobus	170	7	2.1	490	6	O866N2	O866N2 cryctolagus
98	7	2.1	334	2	P77472	P77472 escherichia	171	7	2.1	490	11	Q924K4	Q924K4 mus musculu
99	7	2.1	334	2	P77600	P77600 escherichia	172	7	2.1	491	6	O863A3	O863A3 macaca fasc
100	7	2.1	334	2	O8GGX8	O8GGX8 escherichia	173	7	2.1	492	13	Q7SXV4	Q7SXV4 brachydanio
101	7	2.1	334	2	O8GGX3	O8GGX3 escherichia	174	7	2.1	496	10	O8SSN3	O8SSN3 oryza sativ
102	7	2.1	334	2	O84FH4	O84FH4 acinetobact	175	7	2.1	499	2	O842J1	O842J1 caedibacter
103	7	2.1	334	2	O84D08	O84D08 serratia ma	176	7	2.1	500	5	O95VY0	O95VY0 drosophila
104	7	2.1	334	2	O84D05	O84D05 serratia ma	177	7	2.1	500	5	O9VZU2	O9VZU2 drosophila
105	7	2.1	334	2	O84H92	O84H92 klebsiella	178	7	2.1	502	5	O9XWG6	O9XWG6 caenorhabdi
106	7	2.1	334	2	Q7WUV2	Q7WUV2 klebsiella	179	7	2.1	511	5	O95QI5	O95QI5 caenorhabdi
107	7	2.1	334	16	Q935I6	Q935I6 salmonella	180	7	2.1	520	10	O9FZL2	O9FZL2 arabidopsis
108	7	2.1	334	16	Q934H6	Q934H6 salmonella	181	7	2.1	521	16	O9KND4	O9KND4 vibrio chol
109	7	2.1	334	17	Q97X08	Q97X08 sulfolobus	182	7	2.1	522	4	Q8N6T0	Q8N6T0 homo sapien
110	7	2.1	335	16	O98AA6	O98AA6 rhizobium 1	183	7	2.1	523	16	O8XE85	O8XE85 escherichia
111	7	2.1	335	16	O831E0	O831E0 enterococu	184	7	2.1	530	10	O8LQ57	O8LQ57 oryza sativ
112	7	2.1	337	10	Q7XIQ4	Q7XIQ4 oryza sativ	185	7	2.1	534	10	Q9LGI6	Q9LGI6 oryza sativ
113	7	2.1	339	8	O85Q45	O85Q45 candida gla	186	7	2.1	549	12	O9DWS3	O9DWS3 rat cytomeg
114	7	2.1	344	16	Q97L54	Q97L54 clostridium	187	7	2.1	551	5	O00781	O00781 leishmania
115	7	2.1	345	2	Q9F0B0	Q9F0B0 pseudomonas	188	7	2.1	554	10	O23691	O23691 arabidopsis
116	7	2.1	345	2	Q52481	Q52481 pseudomonas	189	7	2.1	558	5	O20274	O20274 caenorhabdi
117	7	2.1	345	10	O81L9L9	O81L9L9 arabidopsis	190	7	2.1	567	11	O8BK99	O8BK99 mus musculu
118	7	2.1	353	11	Q7TN44	Q7TN44 rattus norv	191	7	2.1	575	10	O949V9	O949V9 arabidopsis
119	7	2.1	357	17	Q9IET3	Q9IET3 aeropyrum p	192	7	2.1	575	10	O9SK52	O9SK52 arabidopsis
120	7	2.1	358	16	O8FEV4	O8FEV4 brucella su	193	7	2.1	587	5	O21893	O21893 caenorhabdi
121	7	2.1	359	2	O8RWM1	O8RWM1 azospirilli	194	7	2.1	593	10	Q7X8T7	Q7X8T7 oryza sativ
122	7	2.1	359	5	O17497	O17497 branchiosto	195	7	2.1	606	8	Q8M434	Q8M434 panthera le
123	7	2.1	359	16	O8YCY1	O8YCY1 brucella me	196	7	2.1	606	17	O973C0	O973C0 sulfolobus
124	7	2.1	355	16	O9XS99	O9XS99 vibrio chol	197	7	2.1	611	13	Q9DDA3	Q9DDA3 xenopus lae
125	7	2.1	365	16	O8D999	O8D999 vibrio vuln	198	7	2.1	614	10	Q9CS11	Q9CS11 arabidopsis
126	7	2.1	368	10	O8VZV7	O8VZV7 arabidopsis	199	7	2.1	624	2	O9RMCA	O9RMCA acinetobact
127	7	2.1	368	10	O9SGW5	O9SGW5 arabidopsis	200	7	2.1	625	10	O93ZH1	O93ZH1 arabidopsis
128	7	2.1	368	16	O87P96	O87P96 vibrio para	201	7	2.1	631	16	O98AT0	O98AT0 rhizobium 1
129	7	2.1	369	2	O9RDY8	O9RDY8 legionella	202	7	2.1	634	5	O8I820	O8I820 bodo ealien
130	7	2.1	369	6	O7YQJ9	O7YQJ9 cryctolagus	203	7	2.1	643	2	O8KJ66	O8KJ66 rhizobium 1
131	7	2.1	375	4	O8WVN8	O8WVN8 homo sapien	204	7	2.1	647	16	Q7V9L8	Q7V9L8 prochloroco
132	7	2.1	377	10	O8LAH2	O8LAH2 arabidopsis	205	7	2.1	649	10	O9LFX1	O9LFX1 arabidopsis
133	7	2.1	377	10	O9W7Y3	O9W7Y3 arabidopsis	206	7	2.1	653	16	O8KFR3	O8KFR3 chlorobium
134	7	2.1	378	11	O8K2Z8	O8K2Z8 mus musculu	207	7	2.1	656	16	O7V430	O7V430 prochloroco
135	7	2.1	378	11	O8BUN2	O8BUN2 mus musculu	208	7	2.1	658	11	O80V96	O80V96 mus musculu
136	7	2.1	379	5	O19391	O19391 caenorhabdi	209	7	2.1	662	5	Q9VDX2	Q9VDX2 drosophila
137	7	2.1	381	16	O8EYI3	O8EYI3 shewanella	210	7	2.1	663	5	O8MSG8	O8MSG8 drosophila
138	7	2.1	386	5	O9BRP3	O9BRP3 caenorhabdi	211	7	2.1	673	11	O8CCN1	O8CCN1 mus musculu
139	7	2.1	387	16	O8EY27	O8EY27 leptospira	212	7	2.1	731	10	Q9ZUE0	Q9ZUE0 arabidopsis
140	7	2.1	390	2	O9JRN4	O9JRN4 actinobacil	213	7	2.1	732	3	Q876G2	Q876G2 saccharomyc
141	7	2.1	396	16	O8CX72	O8CX72 oceanobacil	214	7	2.1	734	3	O8WZU0	O8WZU0 neurospora
142	7	2.1	399	2	O9FLL6	O9FLL6 thermosynec	215	7	2.1	758	16	O9KQC3	O9KQC3 vibrio chol
143	7	2.1	399	16	O8DGB0	O8DGB0 synecococc	216	7	2.1	777	4	Q9BSF5	Q9BSF5 homo sapien
144	7	2.1	400	16	O92VE7	O92VE7 rhizobium m	217	7	2.1	780	5	O8MXZ9	O8MXZ9 tetrahymena
145	7	2.1	402	11	O8BLU3	O8BLU3 mus musculu	218	7	2.1	780	5	O8MQL1	O8MQL1 tetrahymena
146	7	2.1	402	16	Q7UWG3	Q7UWG3 rhodopirell	219	7	2.1	786	16	O8YNK6	O8YNK6 anabaena sp
147	7	2.1	407	2	O9XAY2	O9XAY2 prevotella	220	7	2.1	791	17	Q8TT76	Q8TT76 methanosarc
148	7	2.1	407	8	O8M431	O8M431 panthera ti	221	7	2.1	818	16	O8UCC3	O8UCC3 agrobacteri
149	7	2.1	407	8	O8M432	O8M432 panthera le	222	7	2.1	831	16	O86933	O86933 aquifex aeo
150	7	2.1	422	4	Q7ZTE8	Q7ZTE8 homo sapien	223	7	2.1	846	16	O8ZE91	O8ZE91 yersinia pe
151	7	2.1	422	11	Q7TSS2	Q7TSS2 mus musculu	224	7	2.1	846	16	O8D0H5	O8D0H5 yersinia pe
152	7	2.1	426	10	O9ATH0	O9ATH0 raphanus sa	225	7	2.1	866	16	O8XIW9	O8XIW9 clostridium
153	7	2.1	436	16	O97J44	O97J44 clostridium	226	7	2.1	875	4	Q9H706	Q9H706 homo sapien
154	7	2.1	440	16	O98QF3	O98QF3 mycoplasma	227	7	2.1	877	4	O8ND03	O8ND03 homo sapien
155	7	2.1	442	16	Q7VIN8	Q7VIN8 helicobacte	228	7	2.1	887	16	O883V3	O883V3 pseudomonas
156	7	2.1	443	5	O8T2G0	O8T2G0 dictyosteli	229	7	2.1	902	2	O07686	O07686 listeria se
157	7	2.1	444	2	O56568	O56568 vibrio angu	230	7	2.1	916	3	O875W4	O875W4 saccharomyc
158	7	2.1	444	13	Q7SXY7	Q7SXY7 brachydanio	231	7	2.1	922	10	O9FSE5	O9FSE5 oryza sativ
159	7	2.1	446	10	O8RWK6	O8RWK6 arabidopsis	232	7	2.1	1053	16	O88AL5	O88AL5 pseudomonas
160	7	2.1	452	16	O8FT04	O8FT04 escherichia	233	7	2.1	1055	16	O9A9P9	O9A9P9 caulobacter
161	7	2.1	453	10	O84RN6	O84RN6 oryza sativ	234	7	2.1	1094	4	O8TAB3	O8TAB3 homo sapien
162	7	2.1	453	10	Q7XUF9	Q7XUF9 oryza sativ	235	7	2.1	1115	5	Q27764	Q27764 plasmodium



```
Db 1 MPLFSKSHKNPAEIVKILKONLAILEKQDKTKDASEEVSLSQAMKEILCGTN 54
|||||
RESULT 3
Q8K312 PRELIMINARY; PRT; 205 AA.
AC Q8K312; 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to calcium binding protein, 39 kDa (fragment).
GN CAB39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC029053; AAH29053.1; -.
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
DR NON_TER 1
FT
SQ SEQUENCE 205 AA; 24582 MW; 015261A02F808169 CRC64;

Query Match 6.8%; Score 23; DB 11; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.5e-15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYISKPENLKLMMNLLRDKS 266
Db 109 IMTKYISKPENLKLMMNLLRDKS 131

RESULT 4
Q7Z4X0 PRELIMINARY; PRT; 341 AA.
AC Q7Z4X0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MO25-like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou Y., Yu L., Zhao S.Y.;
RT "Cloning of a new human cDNA homologous to MO25 gene.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134480; AAP97257.1; -.
SQ SEQUENCE 341 AA; 39774 MW; D3A60160E78C7A6D CRC64;

Query Match 6.8%; Score 23; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 5.5e-15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYISKPENLKLMMNLLRDKS 266
Db 245 IMTKYISKPENLKLMMNLLRDKS 267

RESULT 5
Q8VDZ8 PRELIMINARY; PRT; 341 AA.
AC Q8VDZ8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MO25 protein.
GN CAB39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020041; AAH20041.1; -.
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39843 MW; E7FECA529D6FE811 CRC64;

Query Match 6.8%; Score 23; DB 11; Length 341;
Best Local Similarity 100.0%; Pred. No. 5.5e-15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYISKPENLKLMMNLLRDKS 266
Db 245 IMTKYISKPENLKLMMNLLRDKS 267

RESULT 6
Q803V8 PRELIMINARY; PRT; 343 AA.
AC Q803V8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to MO25 protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044172; AAH44172.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 343 AA; 39820 MW; F10450DA0446268A CRC64;

Query Match 5.0%; Score 17; DB 13; Length 343;
Best Local Similarity 100.0%; Pred. No. 9.7e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 219 SENYVTKRQSLKLGEL 235
Db 221 SENYVTKRQSLKLGEL 237

RESULT 7
Q7Z2A5 PRELIMINARY; PRT; 377 AA.
AC Q7Z2A5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein R02E12.2.
GN R02E12.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
```

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA Wilson R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Leimbach D.;  
RT "The sequence of C. elegans cosmid R02E12.";  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U53337; AAP40522.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 377 AA; 43470 MW; 8F9B8AA070216176 CRC64;

Query Match 4.2%; Score 14; DB 5; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPVEY 118  
Db 119 LRRQIGTRSPVEY 132

RESULT 8  
Q21643  
ID Q21643 PRELIMINARY; PRT; 636 AA.  
AC Q21643;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN R02E12.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Leimbach D.;  
RT "The sequence of C. elegans cosmid R02E12.";  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U53337; AAA96186.2; -.  
DR PIR; T16650; T16650.  
DR PIR; T16651; T16651.  
DR WormPep; R02E12.2a; CE28410.

DR InterPro; IPR008938; ARM.  
DR InterPro; IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 636 AA; 72282 MW; 85D5853E9F0E3193 CRC64;  
Query Match 4.2%; Score 14; DB 5; Length 636;  
Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 105 LRRQIGTRSPVEY 118  
Db 378 LRRQIGTRSPVEY 391

RESULT 9  
Q873K5  
ID Q873K5 PRELIMINARY; PRT; 370 AA.  
AC Q873K5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Probable protein required for conidiophore development.  
GN B2314.090.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U.; Aign V.; Hoheisel J.; Brandt P.; Fartmann B.; Holland R.;  
RA Nyakatura G.; Mewes H.W.; Mannhaupt G.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BX284746; CAD70300.1; -.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
SQ SEQUENCE 370 AA; 42812 MW; 2E2C63BD06D8B9E CRC64;

Query Match 2.7%; Score 9; DB 3; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 FHVFKVFVA 282  
Db 303 FHVFKVFVA 311

RESULT 10  
Q8DFL9  
ID Q8DFL9 PRELIMINARY; PRT; 251 AA.  
AC Q8DFL9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Predicted permealase.  
GN VV10192.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H.; Kim S.Y.; Chung S.S.; Kim J.J.; Moon Y.H.; Jeong H.;  
RA Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AR016797; AAO08729.1; -.

DR InterPro; IPR002781; DUF81.  
DR Pfam; PF01925; DUF81; 1.  
KW Complete proteome.  
SQ SEQUENCE 251 AA; 26763 MW; B023D2C45EC6D0C2 CRC64;  
  
Query Match 2.4%; Score 8; DB 16; Length 251;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 75 SSGLLVTL 82  
Db 156 SSGLLVTL 163  
|||||  
|||  
  
RESULT 11  
ID Q8MZ26 PRELIMINARY; PRT; 339 AA.  
AC Q8MZ26; 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DE RE07958p.  
GN CG16912.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY13395; ANM29400.1; -;  
DR FlyBase; FBgn0035064; CG16912.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.  
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.  
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR002307; tRNA-synt\_1.  
DR Pfam; PF00579; tRNA-synt\_1b.  
DR PRINTS; PR01040; TRNASYNTHYR.  
DR TIGRFAMs; TIGR00234; tYRS; 1.  
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
SQ SEQUENCE 339 AA; 37776 MW; DF58AF0E34601FD3D CRC64;  
  
Query Match 2.4%; Score 8; DB 5; Length 339;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 84 ADQLIDF 91  
Db 166 ADQLIDF 173  
|||||  
|||  
  
RESULT 12  
ID Q95S33 PRELIMINARY; PRT; 442 AA.  
AC Q95S33;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DE GM10508p.  
GN CG16912.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY060975; AAL28523.1; -;  
DR FlyBase; FBgn0035064; CG16912.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.  
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002307; Tyr tRNA-synt\_1b.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR PRINTS; PR01040; TRNASYNTHYR.  
DR TIGRFAMs; TIGR00234; tYRS; 1.  
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
SQ SEQUENCE 442 AA; 49898 MW; 347257A9372FB8CC CRC64;  
  
Query Match 2.4%; Score 8; DB 5; Length 442;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 84 ADQLIDF 91  
Db 144 ADQLIDF 151  
|||||  
|||  
  
RESULT 13  
ID Q9W107 PRELIMINARY; PRT; 464 AA.  
AC Q9W107;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CG16912 protein.  
GN CG16912.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,



RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003465; AAF4721.1; -.  
 DR HSP; P00952; 4TSL.  
 DR FlyBase; FBgn0035064; CG16912.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.  
 DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; trna-synt\_1b.  
 DR InterPro; IPR001412; trna-synt\_1.  
 DR InterPro; IPR002307; Tyr-tRNA-synt\_1.  
 DR Pfam; PF00579; trna-synt\_1b; 1.  
 DR PRINTS; PRO1040; TRNASYNTHTYR.  
 DR TIGRFAMs; TIGR00234; tYrs; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 SQ SEQUENCE 464 AA; 52574 MW; E1FFC8F164148C73 CRC64;  
  
 Query Match 2.4%; Score 8; DB 5; Length 464;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 84 ADLQIDF 91  
 Db 166 ADLQIDF 173  
  
 RESULT 14  
 Q87TUI PRELIMINARY; PRT; 576 AA.  
 AC Q87TUI  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Polychyl-phosphate-mannose-protein mannosyltransferase family  
 DE protein.  
 GN PSPT05588.  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,  
 RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,  
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,  
 RA White O., Fraser C., Collier A.,  
 RA "Complete sequence of *Pseudomonas syringae*,"  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE016876; AA059003.1; -.  
 DR TIGR; PSPT05588; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0000030; F:mannosyltransferase activity; IEA.  
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
 DR GO; GO:0006493; P:O-linked glycosylation; IEA.  
 DR InterPro; IPR003342; Glyco\_trans\_39.  
 DR Pfam; PF02366; PWT; 1.  
 KW Transferase; Glycosyltransferase; Complete proteome.  
  
 SQ SEQUENCE 576 AA; 65151 MW; 871FC7A0E8316D47 CRC64;  
  
 Query Match 2.4%; Score 8; DB 16; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 77 GLUVTLLIA 84  
 Db 285 GLUVTLLIA 292  
  
 RESULT 15  
 Q9P8T6 PRELIMINARY; PRT; 695 AA.  
 AC Q9P8T6  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Heat shock protein 80.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20561721; PubMed=11109485;  
 RA Girwitz T.L., Quimet P.M., Kapoor M.;  
 RT "Heat shock protein 80 of *Neurospora crassa*: sequence analysis of the  
 gene and expression during the asexual phase.";  
 RL Can. J. Microbiol. 46:981-991(2000).  
 DR EMBL; AF212996; AAF34607.1; -.  
 DR HSP; P07900; IVER.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003754; F:chaperone activity; IEA.  
 DR InterPro; IPR003594; AtPbind\_ArPase.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00183; HSP90\_1.  
 DR PRINTS; PR00775; HEATSHOCK90.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR PROSITE; PS00298; HSP90; 1.  
 SQ SEQUENCE 695 AA; 78909 MW; 8EAB594ADBA2BC18 CRC64;  
  
 Query Match 2.4%; Score 8; DB 3; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 153 EPLAKIIL 160  
 Db 168 EPLAKIIL 175  
  
 RESULT 16  
 Q76750 PRELIMINARY; PRT; 747 AA.  
 AC Q76750  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative zinc metalloproteinase (Fragment).  
 GN MEP2.  
 OS Haemophilus contortus (Barber pole worm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.  
 OX NCBI\_TaxID=6289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Moredu.  
 RA Redmond D.L., Knox D.P.;  
 RT "Putative metalloproteinase, Mep2, isolated from *Haemonchus*  
 contortus.";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF080117; AAC28740.1; -.

DR GO: GO:0005622; C:intracellular; IEA.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0008237; F:metalloproteinase activity; IEA.  
 DR GO: GO:0004245; F:neprilysin activity; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.  
 DR GO: GO:0006355; F:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR000005; HTHARAC.  
 DR InterPro: IPR000718; Peptidase\_M13.  
 DR InterPro: IPR008753; Peptidase\_M13\_N.  
 DR Pfam: PF01431; Peptidase\_M13; 1.  
 DR Pfam: PF05649; Peptidase\_M13\_N; 1.  
 DR PRINTS: PR00786; NEPRILYSIN.  
 DR PROSITE: PS00041; HTH\_ARAC\_FAMILY\_1; 1.  
 FT NON\_TER 1 747  
 FT TER 747  
 SQ SEQUENCE 747 AA; 84782 MW; DFBEF0FD5608234B CRC64;  
 Query Match 2.4%; Score 8; DB 5; Length 747;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 16 EIVKILKD 23  
 Db 155 EIVKILKD 162  
 RESULT 17  
 Q9V4R2 PRELIMINARY; PRT; 923 AA.  
 ID Q9V4R2;  
 AC Q9V4R2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TOR protein (RE45094p).  
 GN TOR OR CG1389.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Baldwin D.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davesport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,  
 RA Durbin K.J., Evangelista A.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Nurny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dreshek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003840; AAF59203.1; -;  
 DR EMBL: AY071403; AAL49025.1; -;  
 DR HSSP: P11362; 1FGK.  
 DR FlyBase: FBgn0003733; tor.  
 DR GO: GO:0005896; C:plasma membrane; NAS.  
 DR GO: GO:0030381; P:eggshell pattern formation (sensu Insecta); IMP.  
 DR GO: GO:0007362; P:terminal region determination; IMP.  
 DR InterPro: IPR003961; FN III.  
 DR InterPro: IPR000719; Tyr\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR008266; Tyr\_kinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 2.  
 DR SMART: SMO0060; FN3; 1.  
 DR SMART: SMO0219; TyrKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; transferase; tyrosine-protein kinase.  
 SQ SEQUENCE 923 AA; 105200 MW; 746F46E1A4277ACF CRC64;  
 Query Match 2.4%; Score 8; DB 5; Length 923;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 75 SSGLLVTL 82  
 Db 449 SSGLLVTL 456  
 RESULT 18  
 Q8SR14 PRELIMINARY; PRT; 2410 AA.  
 ID Q8SR14;  
 AC Q8SR14;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ubiquitin ligase.  
 GN EC010\_1380.  
 OS Eucephalitozoon cuniculi.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Eucephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-MI;  
 RA Genoscope;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB-MI;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,  
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,

```

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL: AL590449; CAD25857.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro: IPR000569; HECT_domain.
DR Pfam: PF00632; HECT. 1.
DR SMART: SM00119; HECTc. 1.
DR PROSITE: PS00237; HECT. 1.
SQ SEQUENCE 2410 AA; 277308 MW; EF095D412FD6FE7A CRC64;

Query Match 2.4%; Score 8; DB 5; Length 2410;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 PTVEYISA 121
Db 491 PTVEYISA 498
|||||

RESULT 19
ID Q8VHE6 PRELIMINARY; PRT; 4621 AA.
AC Q8VHE6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Axonemal dynein heavy chain 5.
GN DNHC5 OR DNHC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6xCBA/J;
RA Ibanez-Tallon I., Gorokhova S., Heintz N.;
RT "Loss of function of axonemal dynein Dnha5 causes primary ciliary
RT dyskinesia and hydrocephalus.";
RL Hum. Mol. Genet. 11:0-6(2002).
DR EMBL: AF466704; AAL69993.1; -.
DR MGD; MGI:107718; Dnha5.
DR GO: GO:0005858; C:axonemal dynein complex; IMP.
DR GO: GO:0003775; F:axonemal motor activity; IMP.
DR GO: GO:0001539; P:ciliary/flagellar motility; IMP.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR004273; Dynein_heavy.
DR Pfam: PF03028; Dynein_heavy.
DR SMART: SM00382; AAA. 3.
DR PROSITE: PS00639; TH1OL_PROTEASE_HIS; 1.
KW ATP-binding.
SQ SEQUENCE 4621 AA; 527491 MW; E4E044A2D4998298 CRC64;

Query Match 2.4%; Score 8; DB 11; Length 4621;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSSF 308
Db 1815 LIEFLSSF 1822
|||||

RESULT 20
ID Q8MZQ7 PRELIMINARY; PRT; 52 AA.
AC Q8MZQ7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome P450 (Fragment).
GN CYP4C3.
OS Drosophila pachea.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=103846;
RN [1]
RP SEQUENCE FROM N.A.
RA Azadan R.J., Danielson P.B., Fogleman J.C.;
RT "Isolation of a novel cytochrome P450 from Drosophila pachea.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF506532; AAM27400.1; -.
DR FlyBase: FBgn0062760; Dpac\Cyp4c3.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
FT NON_TER 1
SQ SEQUENCE 52 AA; 6072 MW; 98DE3D761CEP7DD6 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 52;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LIGELIL 237
Db 32 LIGELIL 38
|||||

RESULT 21
ID Q94K21 PRELIMINARY; PRT; 56 AA.
AC Q94K21;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF370464; AAK43841.1; -.
SQ SEQUENCE 66 AA; 7464 MW; D51DBE09207D8507 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
Db 2 LLVTLIA 8
|||||

RESULT 22
ID Q81U84 PRELIMINARY; PRT; 93 AA.
AC Q81U84;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN BAI000.
OS Bacillus anthracis (strain Ames).

```

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=198094;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608414; PubMed=12721629;  
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,  
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
RA Holtzapple E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,  
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,  
RA DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Weidner D.,  
RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidner J.F.,  
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,  
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,  
RA Fraser C.M.;  
RT "The genome sequence of *Bacillus anthracis* Ames and comparison to  
RT closely related bacteria."  
RL Nature 423:81-86 (2003).  
DR EMBL; AE017027; AAP24987.1; -.  
DR TIGR; BA1000; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 93 AA; 11272 MW; 131D15787B957F1D CRC64;  
  
Query Match 2.1%; Score 7; DB 16; Length 93;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 IVKILKD 23  
Db 28 IVKILKD 34  
|||||  
|||||  
  
RESULT 23  
Q81H15  
ID Q81H15 PRELIMINARY; PRT; 93 AA.  
AC Q81H15;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN BC1011.  
OS *Bacillus cereus* (strain ATCC 14579 / DSM 31).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=226900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608415; PubMed=12721630;  
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
RA Kapural V., Bhattacharyya A., Reznik G., Mikhalova N., Lapidus A.,  
RA Chu L., Mazur M., Goldsman E., Larsen N., D'Souza M., Walunas T.,  
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,  
RA Overbeek R., Kyrpides N.;  
RT "Genome sequence of *Bacillus cereus* and comparative analysis with  
RT *Bacillus anthracis*."  
RL Nature 423:87-91 (2003).  
DR EMBL; AE017001; AAP07998.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 93 AA; 11095 MW; 82BD67F8587CEE CRC64;  
  
Query Match 2.1%; Score 7; DB 16; Length 93;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 IVKILKD 23  
Db 28 IVKILKD 34  
|||||  
|||||  
  
RESULT 24  
Q88AQ4  
ID Q88AQ4 PRELIMINARY; PRT; 95 AA.  
AC Q88AQ4;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DE Hypothetical protein.  
GN PSPT00332.  
OS *Pseudomonas syringae* (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=DC3000;  
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,  
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,  
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,  
RA White O., Fraser C., Collmer A.;  
RT "Complete sequence of *Pseudomonas syringae*."  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016857; AAO53877.1; -.  
DR TIGR; PSPT00332; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 95 AA; 10614 MW; 131F36C7DA74252 CRC64;  
  
Query Match 2.1%; Score 7; DB 16; Length 95;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 78 LLVTLIA 84  
Db 26 LLVTLIA 32  
|||||  
|||||  
  
RESULT 25  
Q8J2Y6  
ID Q8J2Y6 PRELIMINARY; PRT; 99 AA.  
AC Q8J2Y6;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE RNase.  
GN ODSN.  
OS *Sulfolobus acidocaldarius*.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=2285;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=DSM639;  
RA Schmidt C.L., Henninger T., Schaefer G.;  
RT "A novel RNase from the hyper-thermoacidophilic crenarchaeon  
RT *Sulfolobus acidocaldarius*."  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF202955; AAO13707.1; -.  
DR InterPro; IPR007138; ABM.  
DR Pfam; PF03992; ABM; 1.  
SQ SEQUENCE 99 AA; 11763 MW; EC0D7CD66C464F94 CRC64;  
  
Query Match 2.1%; Score 7; DB 1; Length 99;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 16 EIVKILK 22  
Db 26 EIVKILK 32  
|||||  
|||||  
  
RESULT 26  
Q49831  
ID Q49831 PRELIMINARY; PRT; 103 AA.  
AC Q49831;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B2168_C1.182.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U00018; AA017261.1; -.
DR PIR; S72925; S72925.
SQ SEQUENCE 103 AA; 11211 MW; DID4D669AB82A50F CRC64;

Query Match      2.1%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKIL 21
DB 87 AEIVKIL 93

RESULT 27
Q811K2 PRELIMINARY; PRT; 106 AA.
AC Q811K2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE COG-1B (C.elegans COG-1 protein) (Corresponding sequence
DE R03C1.3b).
DE R03C1.3b).
GN R03C1.3 OR COG-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22370693; PubMed=12482710;
RA Palmer R.E.; Inoue T.; Sherwood D.R.; Jiang L.I.; Sternberg P.W.;
RT "Caenorhabditis elegans cog-1 locus encodes GTX/NKx6.1 homeodomain
RT proteins and regulates multiple aspects of reproductive system
RT development.";
RL Dev. Biol. 252:202-213 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY151210; AA072827.1; -.
DR EMBL; Z82279; CAD62039.1; -.
DR WormPep; R03C1.3b; CE33290.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambrpressor.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
SQ SEQUENCE 106 AA; 12098 MW; EAFBD6B4678F3A20 CRC64;

Query Match      2.1%; Score 7; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQEL 73
DB 50 AQLAQEL 56

RESULT 28
Q42147 PRELIMINARY; PRT; 108 AA.
AC Q42147;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Auxin-responsive GH3 product (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Philipps G.; Gigot C.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z26804; CAA81417.1; -.
DR InterPro; IPR004993; GH3.
DR Pfam; PF03321; GH3; 1.
FT NON TER 1
FT NON TER 108
SQ SEQUENCE 108 AA; 12003 MW; A0D4A0B5EF6E548A CRC64;

Query Match      2.1%; Score 7; DB 10; Length 108;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKQK 298
DB 13 EILLKQK 19

RESULT 29
Q86PF7 PRELIMINARY; PRT; 109 AA.
AC Q86PF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TCC31.39.
GN TCC31.39.
OS Trypanosoma cruzi.
OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brener;
RA Andersson B.; Bontempi E.J.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC137988; AAC44096.1; -.
SQ SEQUENCE 109 AA; 12625 MW; 366F63960B464239 CRC64;

Query Match      2.1%; Score 7; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 MKEILCG 55
DB 49 MKEILCG 55

```

Db 99 MKEILCG 105

## RESULT 30

Q8NY02 PRELIMINARY; PRT; 115 AA.  
AC Q8NY02;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein MW0441.  
GN MW0441.  
OS Staphylococcus aureus (strain MW2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=196620;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiramatsu K.;  
RT "Genome and virulence determinants of high virulence community-  
acquired MRSA.";  
RL Lancet 359:1819-1827 (2002).  
DR EMBL; AP004823; BAB94306.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 115 AA; 12926 MW; CBBDS62267FA7F07 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KDNLAAIL 28

Db 78 KDNLAAIL 84

## RESULT 31

Q899K1 PRELIMINARY; PRT; 116 AA.  
AC Q899K1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Conserved protein.  
GN CTC00172.  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Massachusetts / E88;  
RX MEDLINE=22457253; PubMed=12552129;  
RA Brueggemann H., Baeumer S., Fricke W.F., Wierze A., Liesegang H.,  
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,  
RA Gottschalk G.;  
RT "The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321 (2003).  
DR EMBL; AB015936; AAC04823.1; -;  
KW Complete proteome.  
SQ SEQUENCE 116 AA; 13707 MW; 7FCSAAE44641CAE7 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 KILFSN 163

Db 57 KILFSN 63

## RESULT 32

Q99WB7 PRELIMINARY; PRT; 122 AA.  
AC Q99WB7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein SAV0486.  
GN SAV0486 OR SA0444.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and  
OS Staphylococcus aureus (strain N315).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878, 158879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";  
RL Lancet 357:1225-1240 (2001).  
DR EMBL; AP003359; BAB56648.1; -;  
DR EMBL; AP003130; BAB41674.1; -;  
DR PIR; G89814; G89814.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 122 AA; 13755 MW; 038F597F059A195 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KDNLAAIL 28

Db 85 KDNLAAIL 91

## RESULT 33

Q82QG6 PRELIMINARY; PRT; 133 AA.  
AC Q82QG6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative MerR-family transcriptional regulator.  
GN SAV540.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Onose T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=22608306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial

```

RT microorganism Streptomyces avermitilis.;
RL Nat. Biotechnol. 21:526-531(2003).
RL EMBL; AP005023; BAC68250.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; merR; 1.
DR PRINTS; PR00040; HTHMER.
DR SMART; SM00422; HTH_MER; 1.
DR Complete proteome.
SQ SEQUENCE 133 AA; 14887 MW; FCD73C03FC8BFFCF CRC64;

Query Match 2.1%; Score 7; DB 16; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 VAQLAQE 72
Db 82 VAQLAQE 88
|||||

RESULT 34
Q8YKD8 PRELIMINARY; PRT; 134 AA.
AC Q8YKD8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein All7360.
GN ALL7360.
OS Anabaena sp. (strain PCC 7120).
OG Bacterium; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP MEDLINE=2159285; PubMed=11759840;
RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003603; BAB77344.1; -.
DR PIR; AG2522;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 14894 MW; AA69A2F4E2AEA488 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLAILEK 30
Db 104 NLAILEK 110
|||||

RESULT 35
Q8YKA2 PRELIMINARY; PRT; 134 AA.
AC Q8YKA2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Alr8014.
GN Alr8014.
OS Anabaena sp. (strain PCC 7120).
OG Bacterium; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP MEDLINE=2159285; PubMed=11759840;
RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003601; BAB77118.1; -.
DR PIR; AH2522;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 14913 MW; BOCIDA922114279A CRC64;

Query Match 2.1%; Score 7; DB 16; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLAILEK 30
Db 104 NLAILEK 110
|||||

RESULT 36
Q23309 PRELIMINARY; PRT; 135 AA.
AC Q23309;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN A14G14530.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA Puigdomenech P., Hatzipoulos P., Obermaier B., Duesterhoft A.,
RA Jones J., Palme K., Ansoorge W., Delseny M., Bancroft I., Mewes H.W.,
RA Schueller C., Chalhatzis N.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z97336; CAB10232.1; -.
DR EMBL; AL161539; CAB78495.1; -.
DR PIR; F71407; F71407.
DR Hypothetical protein.
SQ SEQUENCE 135 AA; 15617 MW; A0675960CFC3B311 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 QLIDFEG 93
Db 123 QLIDFEG 129
|||||

RESULT 37
Q88WG2 PRELIMINARY; PRT; 137 AA.
AC Q88WG2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```



```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC
DE 4.2.1.-).
GN FAB22 OR LP_1677.
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIME 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Silezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR ENBL; AL915256; CAB64107.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR006683; Thioester_suff.
DR Pfam; PF03061; 4HBT; 1.
KW Lyase; Complete proteome.
SQ SEQUENCE 137 AA; 15115 MW; 317140F95D1C9D13 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGT 111
Db 108 LRRQIGT 114

RESULT 38
Q9VXD3 PRELIMINARY; PRT; 139 AA.
AC Q9VXD3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE CG5316 protein.
DE CG5316.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196036; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

```

---

```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage I., Morley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR ENBL; AE003726; AAF5667.1; -.
DR FlyBase; FBgn0038704; CGS316.
DR InterPro; IPR001310; HIT.
DR Pfam; PF01230; HIT; 1.
SQ SEQUENCE 139 AA; 16018 MW; 4641400E7B2B5BFD CRC64;

Query Match 2.1%; Score 7; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ISKPENL 255
Db 11 ISKPENL 17

RESULT 39
Q9AJD5 PRELIMINARY; PRT; 140 AA.
AC Q9AJD5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Harpin.
DE Harpin.
GN HRPZ.
OS Pseudomonas syringae (pv. tabaci).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6605;
RA Taguchi F., Tanaka R., Kinoshita S., Ichinose Y., Imura Y., Andi S.,
RA Toyoda K., Shiraishi T., Yamada T.;
RT "HarpinBta from Pseudomonas syringae pv. tabaci is defective and
RT deficient in its expression and HR-inducing activity.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049570; BAB40655.1; -.
DR InterPro; IPR006961; HrpZ.
DR Pfam; PF04877; HrpZ; 1.
SQ SEQUENCE 140 AA; 14801 MW; 713A26C4FA7C08B7 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQL 73
Db 42 AQLAQL 48

RESULT 40
Q7XJG8 PRELIMINARY; PRT; 147 AA.
ID Q7XJG8
AC Q7XJG8;

```

DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE DICHOTOMA-like (Fragment).  
 OS Linaria vulgaris (roadflax).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamiales; Lamiales; Plantaginaceae; Antirrhineae; Linaria.  
 CX NCBI\_TaxID=43171;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Guebitz T., Caldwell A., Hudson A.;  
 RT "Rapid molecular evolution of CYCLOIDEA-like genes in Antirrhinum and  
 its relatives."; Mol. Biol. Evol. 0:0-0(2003).  
 RL EMBL: AY316702; AAP84108.1; -.  
 DR NON\_TER 147 147  
 FT NON\_TER 147 147  
 SQ SEQUENCE 147 AA; 16623 MW; 347F7C17D23EA393 CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 218 QSENYVT 224  
 Db 129 QSENYVT 135  
 RESULT 41  
 Q7XJG9 PRELIMINARY; PRT; 151 AA.  
 AC Q7XJG9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE DICHOTOMA-like (Fragment).  
 OS Linaria maroccana.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamiales; Lamiales; Plantaginaceae; Antirrhineae; Linaria.  
 CX NCBI\_TaxID=237459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Guebitz T., Caldwell A., Hudson A.;  
 RT "Rapid molecular evolution of CYCLOIDEA-like genes in Antirrhinum and  
 its relatives."; Mol. Biol. Evol. 0:0-0(2003).  
 RL EMBL: AY316701; AAP84107.1; -.  
 DR NON\_TER 151 151  
 FT NON\_TER 151 151  
 SQ SEQUENCE 151 AA; 16940 MW; DCE4DF01B3BC6318 CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 218 QSENYVT 224  
 Db 133 QSENYVT 139  
 RESULT 42  
 Q92C12 PRELIMINARY; PRT; 154 AA.  
 AC Q92C12;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein lin1379.  
 GN lin1379.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria species.*";  
 RL Science 294:849-852(2001).  
 DR EMBL: AL596168; CAC96610.1; -.  
 DR FIR; AB1605; AB1605.  
 DR Listlist; lin01379; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 154 AA; 18073 MW; 5A95AB680C57F330 CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 102 NNILRQ 108  
 Db 99 NNILRQ 105  
 RESULT 43  
 Q80IV1 PRELIMINARY; PRT; 156 AA.  
 AC Q80IV1;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE L\* protein.  
 OS Theiler-like virus NGS910.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Cardiovirus.  
 CX NCBI\_TaxID=204729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NGS910;  
 RA Ohsawa K., Watanabe Y., Miyata H., Sato H.;  
 RT "Genetic Analysis of a Theiler-Like Virus Isolated from Rats.";  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB090161; BACS8036.1; -.  
 SQ SEQUENCE 156 AA; 18183 MW; 9D8DCEDC0A6E9DE0 CRC64;  
 Query Match 2.1%; Score 7; DB 12; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 45 SIQAMKE 51  
 Db 84 SIQAMKE 90  
 RESULT 44  
 Q7XUY5 PRELIMINARY; PRT; 157 AA.  
 AC Q7XUY5;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE OSJNB0048E02.12 protein.  
 GN OSJNB0048E02.12.  
 OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Meng Q.J., Zhang L., Li T.,  
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL606653; CAD40936.1; -;  
 DR EMBL; AL606653; CAD40936.1; -;  
 SQ SEQUENCE 157 AA; 17266 MW; DA8AC54057F26D3E CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 DKASEEV 42  
 Db 127 DKASEEV 133  
 RESULT 45  
 Q8ABD4  
 ID Q8ABD4 PRELIMINARY; PRT; 162 AA.  
 AC Q8ABD4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN B70176.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Bacteroidaceae; Bacteroides.  
 OX NCBI\_TaxID=818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=12663928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chiang H.C., Hooper L.V., Gordon J.R.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
 RL Science 299:2074-2076(2003).  
 DR EMBL; AE016926; AA075283.1; -;  
 DR EMBL; AE016926; AA075283.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 162 AA; 18772 MW; B9C22B0528F0EF5A CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 310 KERTDDE 316  
 Db 111 KERTDDE 117  
 RESULT 46  
 ID Q8KY28 PRELIMINARY; PRT; 164 AA.  
 AC Q8KY28;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Histidinol-phosphate aminotransferase (Fragment).  
 OS Aminomonas aminovorans.  
 OC Bacteria; Proteobacteria.  
 OX NCBI\_TaxID=135579;  
 Query Match 2.1%; Score 7; DB 16; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=NCIME 11288;  
 RA Taylor E.J., Smith N.L., Colby J., Black G.W.;  
 RT "The 3-hexulose-6-phosphate synthase-encoding gene of the obligate  
 RT methylotroph Aminomonas aminovorans C2A1 is not present on a gene  
 RT cluster encoding other enzymes of the ribulose monophosphate  
 RT pathway."  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF294615; AAM88571.1; -;  
 DR GO; GO:0008483; P:transaminase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR004839; Aminotrans\_I/II.  
 DR InterPro; IPR001917; Aminotrans\_III.  
 DR Pfam; PF00155; aminotran\_1\_2; 1.  
 DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
 KW Transferase; Aminotransferase.  
 FT NON\_TER 1  
 SQ SEQUENCE 164 AA; 17791 MW; 2736F03BB5D1830E CRC64;  
 Query Match 2.1%; Score 7; DB 2; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 81 TLIADLQ 87  
 Db 85 TLIADLQ 91  
 RESULT 47  
 Q931S7  
 ID Q931S7 PRELIMINARY; PRT; 165 AA.  
 AC Q931S7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ABC transporter homolog.  
 GN SAV1318  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2111952; PubMed=11418146;  
 RA Kuroda M., Ohka T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguni A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 RT aureus."  
 RL Lancet 357:1225-1240(2001).  
 DR EMBL; AF003362; BAB57480.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0004009; P:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 165 AA; 18835 MW; 7A488E4DB5A9C540 CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 215 KLIQSEN 221

```
Db          63 KLLQSEN 69
|||||
RESULT 48
Q922L7      PRELIMINARY; PRT; 171 AA.
AC Q922L7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RA0439.
GN RA0439 OR SWA0809.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federapfel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
DR EMBL; AE007234; AAK65097.1; -.
DR FIR; G95316; G95316.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 171 AA; 13006 MW; 2399EE398413CA0 CRC64;

Query Match          2.1%; Score 7; DB 16; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 PTEAVAQ 68
|||||
DB 61 PTEAVAQ 67

RESULT 49
O66635      PRELIMINARY; PRT; 179 AA.
ID O66635;
AC O66635;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcriptional regulator (TETR/ACRR family).
GN ACRR3 OR AQ_281.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VP5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anuj M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358 (1998).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AE000682; AAC06593.1; -.
DR FIR; G70325; G70325.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

```
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr; 1.
DR PRINTS; PR00455; HTHTR.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 179 AA; 20720 MW; 1E56E2FDA4B4C850 CRC64;

Query Match          2.1%; Score 7; DB 16; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIILF 161
|||||
DB 100 LAKIILF 106
|||||

RESULT 50
Q89US1      PRELIMINARY; PRT; 180 AA.
ID Q89US1;
AC Q89US1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Brl1340 protein.
GN Brl1340.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasanoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AF005939; BAC46605.1; -.
KW Complete proteome.
SQ SEQUENCE 180 AA; 18714 MW; 60C7654EAE004943 CRC64;

Query Match          2.1%; Score 7; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
|||||
DB 24 LLVTLIA 30
|||||

RESULT 51
Q8EM80      PRELIMINARY; PRT; 181 AA.
ID Q8EM80;
AC Q8EM80;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE H(+)-transporting ATP synthase delta chain (EC 3.6.1.34).
GN ATPH OR OB2978.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935 (2002).
```

```

DR EMBL; AP004603; BAC14934.1; -.
DR GO; GO:0016787; F-hydrolase activity; IEA.
DR GO; GO:0006754; P:ATP biosynthesis; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR000711; ATPsynth_OSCP.
DR Pfam; PF00213; OSCP; 1.
DR PRINTS; PR00125; ATPASDELTA.
DR TIGRFAMs; TIGR01145; ATP synt delta; 1.
DR PROSITE; PS00389; ATPASE_DELTA; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 181 AA; 20306 MW; 1E81347B683D4AD CRC64;

Query Match          2.1%; Score 7; DB 16; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 197 LVADFLF 203
Db 27 LVADFLF 33

RESULT 52
O55132 PRELIMINARY; PRT; 185 AA.
AC O55132
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SV129;
RX MEDLINE=20016253; PubMed=10547269;
RA Silva A.M., Bottrel R.L.A., Reis L.F.L.;
RT "Identification and partial characterization of FRAG-6, a novel
RT interferon-stimulated gene that is expressed in an IFP-1-independent
RT manner."
RL Cytokine 11:813-821(1999).
DR EMBL; U76754; AAB92383.1; -.
KW Hypothetical protein.
FT NON TER 185
SQ SEQUENCE 185 AA; 21318 MW; 843031PALEB3C54B CRC64;

Query Match          2.1%; Score 7; DB 11; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 LCGTNEK 59
Db 30 LCGTNEK 36

RESULT 53
Q87I75 PRELIMINARY; PRT; 187 AA.
ID Q87I75
AC Q87I75
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN VPA0731.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
CX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;

RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
DR EMBL; AP005086; BAC62074.1; -.
DR InterPro; IPR006664; Bac OmpA.
DR InterPro; IPR006665; OmpA/MotB.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01021; OMPADOMAIN.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 187 AA; 20814 MW; F8A1476C6DE32B17 CRC64;

Query Match          2.1%; Score 7; DB 16; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 QSENYVT 224
Db 23 QSENYVT 29

RESULT 54
Q8YRB9 PRELIMINARY; PRT; 191 AA.
ID Q8YRB9
AC Q8YRB9
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Alr3529.
GN Alr3529.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
CX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003593; BAB75228.1; -.
DR PIR; AB2247; AB2247.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 191 AA; 22465 MW; AE36336359FC90DE CRC64;

Query Match          2.1%; Score 7; DB 16; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 LIADLQL 88
Db 44 LIADLQL 50

RESULT 55
Q8N7X8 PRELIMINARY; PRT; 197 AA.
ID Q8N7X8
AC Q8N7X8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ40235.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK097554; BAC05098.1; -.
DR InterPro; IPR007110; IG-like.
KW Hypothetical protein.
SQ SEQUENCE 197 AA; 21283 MW; 4E6807E266357A25 CRC64;

Query Match      2.1%; Score 7; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 ILEKQDK 33
    |||||
Db 188 ILEKQDK 194

RESULT 56
Q96J08      PRELIMINARY;      PRT;      199 AA.
AC Q96J08;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein-protein ligase (EC 6.3.2.19) (Ubiquitin-
DE conjugating enzyme E2) (Ubiquitin carrier protein)
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; BC006827; AA06827.1; -.
DR GO; GO:0016874; P:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UQ_con; 1_
DR ProDom; PD000461; UEQ_conjugat; 1.
DR PROSITE; PS00127; UBIQUITIN_CONJUGAT 2; 1.
KW Hypothetical protein; Ligase; Ubl conjugation pathway.
FT NON TER 1
SQ SEQUENCE 199 AA; 22836 MW; 0523FA017EBE8820 CRC64;

Query Match      2.1%; Score 7; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLAILEK 30
    |||||
Db 64 NLAILEK 70

```

```

RESULT 57
Q9N8D8      PRELIMINARY;      PRT;      199 AA.
AC Q9N8D8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ39623.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK096942; BAC04908.1; -.
DR InterPro; IPR001279; Blactmase-like.
DR Pfam; PF00753; lactamase_B; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 22330 MW; 90A97FC36A1AD66A CRC64;

Query Match      2.1%; Score 7; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGLL 79
    |||||
Db 57 LYSSGLL 63

RESULT 58
Q981K6      PRELIMINARY;      PRT;      201 AA.
AC Q981K6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein ml19335.
GN ML19335.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003015; BAB54703.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 201 AA; 23755 MW; 4E410B65A36DE7A1 CRC64;

Query Match      2.1%; Score 7; DB 16; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 229 LKLGEL 235
DB 173 LKLGEL 179

RESULT 59
Q96MV4 PRELIMINARY; PRT; 217 AA.
AC Q96MV4 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein FJ331826 (EC 6.3.2.19) (Ubiquitin-conjugating
DE enzyme E2) (Ubiquitin-protein ligase) (Ubiquitin carrier
DE protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -1- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; AK056388; BAB71173.1; -.
DR Genbank; HGNC:15698; UBE2Q.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR006068; Ubq_conjugat.
DR Pfam; PF00179; Ubq_con; 1.
DR Prodom; PD000461; Ubq_conjugat; 1.
DR PROSITE; PS0127; UBIQUITIN_CONJUGAT 2; 1.
KW Hypothetical protein; Ligase; Ub1 conjugation pathway.
SQ SEQUENCE 217 AA; 24277 MW; 3C74AF937AE2BCE CRC64;

Query Match 2.1%; Score 7; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEX 30
DB 21 NLAILEX 27

RESULT 60
Q7XQH7 PRELIMINARY; PRT; 218 AA.
AC Q7XQH7 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE OSJNBa0067K08.19 protein.
DE OSJNBa0067K08.19 protein.
GN OSJNBa0067K08.19.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL606627; CAE03116.1; -.
SQ SEQUENCE 218 AA; 24583 MW; 80783560AFAA416A CRC64;

Query Match 2.1%; Score 7; DB 10; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
DB 142 SLKLGE 148

RESULT 61
O18347 PRELIMINARY; PRT; 221 AA.
AC O18347 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JAN-1998 (TREMBlrel. 24, Last annotation update)
DE Venom allergen antigen 5-like protein.
GN ND1VA833.
OS Dirofilaria immitis (Cane heartworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Dirofilaria.
CX NCBI_TaxID=6287;
RN [1]
RP SEQUENCE FROM N.A.
RA Tripp C.A., Wisniewski N.;
RA "Dirofilaria immitis larval cDNA clone isolated with immune dog
RA sera.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001100; AAB62535.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR001283; Allrgn_V5/Tpx1.
DR Pfam; PF00188; SCP; 1.
DR PRINTS; PR00837; VSTPLXKE.
DR Prodom; PD000542; Allrgn_V5/Tpx1; 1.
DR SMART; SM00198; SCP; 1.
DR PROSITE; PS01010; SCP AG5 PR1 SC7 2; 1.
SQ SEQUENCE 221 AA; 25158 MW; 5542B42B455046F7 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 RDLKTA 336
DB 109 RDLKTA 115

RESULT 62
Q9D614 PRELIMINARY; PRT; 221 AA.
AC Q9D614 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 21, Last annotation update)
DE 2900024010Rik protein.

```



```

GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX MEDLINE=21085660; PubMed=1117851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK013594; BAB28918.1; -.
DR MGD; MGI:1920102; 2900024O10RIK.
DR InterPro; IPR001279; Blackmase-like.
DR Pfam; PF00753; lactamase B; 1.
SQ SEQUENCE 221 AA; 24486 MW; 9A604DFBES64DFE CRC64;

Query Match 2.1%; Score 7; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGILL 79
DB 57 LYSSGILL 63

RESULT 63
Q33S74 ID Q83S74 PRELIMINARY; PRT; 222 AA.
AC Q83S74;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Putative transposase.
GN SF0228.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Ou D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015059; AAN41890.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.

DR Pfam; PF00665; rve; 1.
KW Complete proteome.
SQ SEQUENCE 222 AA; 24899 MW; 5E42A1062243F9EA CRC64;

Query Match 2.1%; Score 7; DB 16; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 MLRECIR 151
DB 1 MLRECIR 7

RESULT 64
Q82P19 ID Q82P19 PRELIMINARY; PRT; 227 AA.
AC Q82P19;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV114.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Kinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RW EMBL; AP005025; BAC68824.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 227 AA; 23392 MW; ED9D02BE1BF9D08A CRC64;

Query Match 2.1%; Score 7; DB 16; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQLA 70
DB 211 EAVAQLA 217

RESULT 65
Q9BVX5 ID Q9BVX5 PRELIMINARY; PRT; 230 AA.
AC Q9BVX5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical ubiquitin-protein ligase (EC 6.3.2.19) (Ubiquitin-
DE conjugating enzyme E2) (Ubiquitin carrier protein)
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN  SEQUENCE FROM N.A.
RC  TISSUE=Cervix;
RA  Strausberg R.;
RL  Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC  PROTEINS (BY SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC  DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC  -1- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC  -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC  THIOLESTER FORMATION (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR  EMBL; BC000848; AA00848.1; -.
DR  GO; GO:0016874; F:ligase activity; IEA.
DR  GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR  GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR  GO; GO:0006512; P:ubiquitin cycle; IEA.
DR  InterPro; IPR000608; UBQ_conjugat.
DR  Pfam; PF00179; UQ_con; 1.
DR  ProDom; PD000461; UBQ_conjugat; 1.
DR  SMART; SM00212; UBCC; 1.
DR  PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW  Hypothetical protein; Ligase; Ub1 conjugation pathway.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 230 AA; 25882 MW; 97C13304B8E1D6E7 CRC64;

Query Match          2.1%; Score 7; DB 4; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  24 NLAILEK 30
Db  34 NLAILEK 40

RESULT 66
Q8K2T0 PRELIMINARY; PRT; 231 AA.
AC Q8K2T0;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE RIKEN cDNA 2310012M18 gene (Fragment).
GN 2310012M18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030044; AA030044.2; -.
DR MGD; MGI:1917343; 2310012M18RIK.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UQ_con; 1.
DR SMART; SM00212; UBCC; 1.
DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 231 AA; 26011 MW; BDFCB7D7F5D51F1A CRC64;

Query Match          2.1%; Score 7; DB 11; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  24 NLAILEK 30
Db  35 NLAILEK 41
```

```

RESULT 67
Q9DVX8 PRELIMINARY; PRT; 251 AA.
AC Q9DVX8;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-WAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE PXORF55 peptide.
GN PXORF55.
OS Plutella xylostella granulovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=98383;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X1;
RX MEDLINE=20455581; PubMed=10998336;
RA Hashimoto Y., Hayakawa T., Ueno Y., Fujita T., Sano Y., Matsumoto T.;
RT "Sequence analysis of the Plutella xylostella granulovirus genome.";
RL Virology 275:358-372(2000).
DR EMBL; AF270937; AG27353.1; -.
DR InterPro; IPR007983; Baculo LEF-1.
DR Pfam; PF05319; Baculo LEF-1; 1.
SQ SEQUENCE 251 AA; 29931 MW; 167FB1736F7B9FB8 CRC64;

Query Match          2.1%; Score 7; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  260 NLLRDKS 266
Db  155 NLLRDKS 161

RESULT 68
Q87RH6 PRELIMINARY; PRT; 251 AA.
AC Q87RH6;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN VP0818.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AF005075; BAC59081.1; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Complete proteome.
SQ SEQUENCE 251 AA; 26482 MW; 2FE09CD409388D45 CRC64;

Query Match          2.1%; Score 7; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  76 SGLLVTL 82
Db  157 SGLLVTL 163

RESULT 69
```

```
Q9D7E1
ID Q9D7E1 PRELIMINARY; PRT; 253 AA.
AC Q9D7E1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310012M18 product:NICE-5 protein (PRO3094) homolog.
DE 2310012M18RIK.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nonura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=21085660; PubMed=11217851;
RX RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=99279253; PubMed=10349636;
RX Carninci P., Hayashizaki Y.;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=20499374; PubMed=11042159;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=20530913; PubMed=11076861;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
```

```
Genome Res. 10:1757-1771(2000).
RL EMBL; AK009324; BAB26217.2; -.
DR MGD; MG1:1917343; 2310012M18RIK.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; Ubq_conjugat.
DR Pfam; PF00179; Ubq_con; 1.
DR ProDom; PD000461; Ubq_conjugat; 1.
DR SMART; SM00212; UBC; 1.
DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
SQ SEQUENCE 253 AA; 28487 MW; 8F9519FF26585CF CRC64;

Query Match 2.1%; Score 7; DB 11; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLAILEK 30
Db 57 NLAILEK 63

RESULT 70
Q8FS91
ID Q8FS91 PRELIMINARY; PRT; 253 AA.
AC Q8FS91
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase.
GN CE0506.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005215; BAC17316.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Complete proteome.
SQ SEQUENCE 253 AA; 27807 MW; 51AB97E3BFBC0CB7 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 YEKLLQS 219
Db 246 YEKLLQS 252

RESULT 71
Q8BVX5
ID Q8BVX5 PRELIMINARY; PRT; 255 AA.
AC Q8BVX5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to NIC-5 protein homolog.
GN 3010021M21RIK.
OS Mus musculus (Mouse).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK076148; BAC36218.1; -  
 DR MGD; MGI:1924088; 3010021M21Rik.  
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBQC; 1.  
 DR PROSITE; PS0127; UBIQUITIN CONJUGAT 2; 1.  
 SQ SEQUENCE 255 AA; 28916 MW; 2F5077B3FAC2D6DF CRC64;  
 Query Match 2.1%; Score 7; DB 11; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 24 NLAILEX 30  
 DB 59 NLAILEX 65  
 RESULT 72  
 O17978  
 ID O17978 PRELIMINARY; PRT; 256 AA.  
 AC O17978;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE R03C1.3 protein (COG-1A).  
 GN R03C1.3 OR COG-1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Steward C.A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBSJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22370693; PubMed=12482710;  
 RA Palmer R.E., Inoue T., Sherwood D.R., Jiang L.I., Sternberg P.W.;  
 RT "Caenorhabditis elegans cog-1 locus encodes GTX/NKx6.1 homeodomain  
 RT proteins and regulates multiple aspects of reproductive system  
 RT development.";  
 RL Dev. Biol. 252:202-213(2002).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; Z83279; CAB05260.2; -  
 DR EMBL; AY151209; AAN72826.1; -  
 DR PIR; T23866; T23866.  
 DR HSSP; P14653; 1B72.  
 DR WormPep; R03C1.3a; CE25964.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambdarepressr.  
 DR Pfam; PF00046; Homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX 1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 256 AA; 28046 MW; E961B929941CC719 CRC64;  
 Query Match 2.1%; Score 7; DB 5; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 67 AQLAQL 73  
 DB 200 AQLAQL 206  
 RESULT 73  
 Q7U4F8  
 ID Q7U4F8 PRELIMINARY; PRT; 258 AA.  
 AC Q7U4F8;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative ABC transporter, multidrug efflux family.  
 GN SYNW2110.  
 OS Synechococcus sp. (strain WH8102).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=84588;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22825697; PubMed=12917641;  
 RA Palenik B., Bråhamsha B., Larimer F.W., Land M., Hauser L., Chain P.,  
 RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,  
 RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;  
 RT "The genome of a motile marine Synechococcus.";  
 RL Nature 424:1037-1042(2003).  
 RL EMBL; BX569694; CAE08625.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 258 AA; 27641 MW; 419DDAF0E8894F3E CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 105 LRRQIGT 111  
 DB 96 LRRQIGT 102  
 RESULT 74  
 Q84RB8  
 ID Q84RB8 PRELIMINARY; PRT; 264 AA.  
 AC Q84RB8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE DICHOTOMA-like protein (Fragment).  
 GN DICH.  
 OS Linaria canadensis.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Lamiaceae; Plantaginaceae; Antirrhineae; Linaria.  
 OX NCBI\_TaxID=69914;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22569055; PubMed=12679544;  
 RA Hileman L.C., Baum D.A.;  
 RT "Why Do Paralogs Persist? Molecular Evolution of CYCLOIDEA and Related

RT Floral Symmetry Genes in Antirrhineae (Veronicaaceae).";  
 RL Mol. Biol. Evol. 20:591-600(2003).  
 DR EMBL; AF512589; AAP03336.1; -.  
 DR InterPro; IPR005333; TCP.  
 DR Pfam; PF03634; TCP; 1.  
 FT NON TER 264 264  
 FT NON TER 264 264  
 SQ SEQUENCE 264 AA; 29428 MW; FEC08C31C452AB5C CRC64;

Query Match 2.1%; Score 7; DB 10; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 218 QSENVVT 224  
 Db 244 QSENVVT 250  
 |||||

## RESULT 75

Q84RB7 PRELIMINARY; PRT; 267 AA.  
 AC Q84RB7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE DICHOTOMA-like protein (Fragment).  
 GN DICH.  
 OS Linaria vulgaris (Toadflax).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Plantaginaceae; Antirrhineae; Linaria.  
 OC NCBI\_TaxID=43171;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22569055; PubMed=12679544;  
 RA Hileman L.C.; Baum D.A.;  
 RT "Why Do Paralogous Persist? Molecular Evolution of CYCLOIDEA and Related  
 FL Floral Symmetry Genes in Antirrhineae (Veronicaaceae).";  
 RL Mol. Biol. Evol. 20:591-600(2003).  
 DR EMBL; AF512590; AAP03337.1; -.  
 DR InterPro; IPR005333; TCP.  
 DR Pfam; PF03634; TCP; 1.  
 FT NON TER 267 267  
 FT NON TER 267 267  
 SQ SEQUENCE 267 AA; 29874 MW; 6DF6F2DBA2DA08CC CRC64;

Query Match 2.1%; Score 7; DB 10; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 218 QSENVVT 224  
 Db 247 QSENVVT 253  
 |||||

## RESULT 76

Q8BKU7 PRELIMINARY; PRT; 272 AA.  
 AC Q8BKU7;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Dynamalin-like 120 kDa protein (Fragment).  
 GN OPAL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RL 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK050383; BAG34224.1; -.  
 DR MGD; MGI:1921393; Opal.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 FT NON TER 272 272  
 FT NON TER 272 272  
 SQ SEQUENCE 272 AA; 31438 MW; 9152E557CFF9A21F3 CRC64;

Query Match 2.1%; Score 7; DB 11; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 231 LLGELIL 237  
 Db 210 LLGELIL 216  
 |||||

## RESULT 77

Q9X793 PRELIMINARY; PRT; 275 AA.  
 AC Q9X793;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Acyl-[ACP] desaturase.  
 GN DSSA2 OR MJ1952 OR MLCB1222.21.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T.; Eigmeier K.; Parkhill J.; James K.D.; Thomson N.R.;  
 RA Wheeler P.R.; Honore N.; Garnier T.; Churcher C.; Harris D.;  
 RA Mungall K.; Basham D.; Brown D.; Chillingworth T.; Connor R.;  
 RA Davies R.M.; Devlin K.; Duthoy S.; Feltwell T.; Fraser A.; Hamlin N.;  
 RA Holroyd S.; Hornsby T.; Jagels K.; Lacroix C.; Maclean J.; Moule S.;  
 RA Murphy L.; Oliver K.; Quail M.A.; Rajandream M.A.; Rutherford K.M.;  
 RA Rutter S.; Seeger K.; Simon S.; Simmonds M.; Skelton J.; Squares R.;  
 RA Squares S.; Stevens K.; Taylor K.; Whitehead S.; Woodward J.R.;  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus.";  
 RL Nature 409:1007-1011(2001).  
 DR EMBL; AL049491; CAB39827.1; -.  
 DR EMBL; AL583923; CAC30907.1; -.  
 DR PIR; C87153; C87153.  
 DR Leproma; MJ1952; -.  
 DR GO; GO:0045300; F:acyl-[acyl-carrier protein] desaturase acti. . .; IEA.  
 DR GO; GO:0006631; P:fatty acid metabolism; IEA.  
 DR InterPro; IPR005067; PA.desat.  
 DR Pfam; PF03405; PA.desaturase\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 275 AA; 31139 MW; 06C1B9F7087DE46C CRC64;

Query Match 2.1%; Score 7; DB 16; Length 275;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 ILKDNLA 26  
 Db 74 ILKDNLA 80  
 |||||

## RESULT 78

OS3442 PRELIMINARY; PRT; 275 AA.  
 AC OS3442;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)



```

RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038230; AAH38230.1; -.
DR InterPro; IPR001279; Blactmase-like.
DR Pfam; PF00753; lactamase_B; 1.
SQ SEQUENCE 279 AA; 31341 MW; 6343E17E62DC492D CRC64;

Query Match      2.1%; Score 7; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGLL 79
Db 57 LYSSGLL 63

RESULT 82
Q8BL86 PRELIMINARY; PRT; 279 AA.
AC
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical metallo-beta-lactamase superfamily containing
DE protein.
GN 2900024O10RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK046032; BAC32577.1; -.
DR MGD; MGI:1920102; 2900024O10RIK.
DR InterPro; IPR001279; Blactmase-like.
DR Pfam; PF00753; lactamase_B; 1.
KW Hypothetical protein.
SQ SEQUENCE 279 AA; 31206 MW; 83F882FF44580DF1 CRC64;

Query Match      2.1%; Score 7; DB 11; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGLL 79
Db 57 LYSSGLL 63

RESULT 83
Q97W91 PRELIMINARY; PRT; 282 AA.
AC
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Predicted methyltransferase.
GN CAC0307.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007545; AAK78288.1; -.
DR PIR; E96937; E96937.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000878; Cor/por Metransf.
DR InterPro; IPR008189; UPF0011.
DR Pfam; PF00590; TP methylase; 1.
DR TIGRFAMs; TIGR00036; TIGR00096; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 282 AA; 31981 MW; EEC11B67310B6C1F CRC64;

Query Match      2.1%; Score 7; DB 16; Length 282;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 226 RQSLKLL 232
Db 39 RQSLKLL 45

RESULT 84
Q8DPF4 PRELIMINARY; PRT; 283 AA.
ID
AC Q8DPF4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter membrane-spanning permease-oligopeptide
DE transport.
GN APC OR SPR1192.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RT J. Bacteriol. 183:5709-5717(2001).
RL EMBL; AE008491; AAK99995.1; -.
DR PIR; G98020; G98020.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR InterPro; IPR000437; ProK lipoprot_s.
DR Pfam; PF00528; BPD transp; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 31176 MW; A08B18B83E18B750 CRC64;

Query Match      2.1%; Score 7; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 IFNNILR 106

```



```
Db 189 IFNNLR 195
|||||
RESULT 85
Q99UF7 PRELIMINARY; PRT; 299 AA.
AC Q99UF7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Sall156 (NM1206 protein).
GN Sall156 OR MW1206.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879, 196820;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NM115;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayaishi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NM2;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827 (2002).
DR EMBL; AP003133; BAB424.1; -.
DR EMBL; AP004826; BAB95071.1; -.
DR PIR; H89906; H89906.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 299 AA; 34577 MW; 451484C431B3AEF0 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 KLLQSEN 221
DB 63 KLLQSEN 69
|||||

RESULT 86
O74388 PRELIMINARY; PRT; 301 AA.
AC O74388;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
```

RA Van Etten J.L.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Van Etten J.L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Graves M.V., Van Etten J.L.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42580; AAC96435.1; -.  
 DR PIR; T17557; T17557.  
 SQ SEQUENCE 309 AA; 33210 MW; 5397BE4D0A235C26 CRC64;

Query Match 2.1%; Score 7; DB 12; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 TFDIASD 182  
 |||||  
 Db 120 TFDIASD 126

## RESULT 88

Q8X5K0  
 ID Q8X5K0 PRELIMINARY; PRT; 309 AA.

AC Q8X5K0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 2-hydroxy-6-ketono-2,4-dienedioic acid hydrolase.  
 GN MHPC OR 20447 OR ECS0404.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=83334;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 Raofai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kunara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL; AE005214; AAC54700.1; -.  
 DR EMBL; AP002551; BAB33827.1; -.  
 DR PIR; D90679; D90679.  
 DR PIR; H85529; H85529.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.

DR InterPro; IPR000073; A/b hydrolase.  
 DR InterPro; IPR003089; AB hydrolase.  
 DR InterPro; IPR000639; Epox\_hydrolase.  
 DR InterPro; IPR000373; Ser\_estrs.  
 DR Pfam; PF00561; abhydrolase; 1.  
 DR PRINTS; PR00111; ABHYDROLASE.  
 DR PRINTS; PR00412; EPOXYHYDROLASE.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 309 AA; 34615 MW; E2B51850912969CE CRC64;

Query Match 2.1%; Score 7; DB 16; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 253 ENLKMM 259  
 |||||  
 Db 189 ENLKMM 195

## RESULT 89

Q7ZA52  
 ID Q7ZA52 PRELIMINARY; PRT; 312 AA.

AC Q7ZA52;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Aldose reductase.  
 OS Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetes; incertae sedis; Magnaportheaceae; Magnaporthe.  
 OX NCBI\_TaxID=148305;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Zheng Y.-J., Tao Y., Zhang W., Jordan D.;  
 RT "Inhibition of fungal aldose reductase";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY033888; AAC55762.1; -.  
 SQ SEQUENCE 312 AA; 34368 MW; 10FE14BC676B5C55 CRC64;

Query Match 2.1%; Score 7; DB 3; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 QSLKLG 233  
 |||||  
 Db 94 QSLKLG 100

## RESULT 90

O87260  
 ID O87260 PRELIMINARY; PRT; 314 AA.

AC O87260;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Lactococcus lactis.  
 OG Plasmid pMRC01.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 OX NCBI\_TaxID=1358;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=DPC3147;  
 RX MEDLINE=99000510; PubMed=9767571;  
 RA Dougherty B.A., Hill C., Weidman J.F., Richardson D.R., Venter J.C.,  
 RA Ross R.P.;

RT "Sequence and analysis of the 60 kb conjugative, bacteriocin-producing  
 plasmid pMRC01 from Lactococcus lactis DPC3147";  
 RL Mol. Microbiol. 29:1029-1038(1998).  
 DR EMBL; AE001272; AAC56050.1; -.  
 DR PIR; T43132; T43132.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; Plasmid.

SQ SEQUENCE 314 AA; 37078 MW; B6D0B01B2065C525 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 RDFPKYV 172  
 |||||  
 Db 182 RDFPKYV 188

RESULT 91

O32778 PRELIMINARY; PRT; 314 AA.

AC O32778;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Lactococcus lactis.  
 OG Plasmid pK214.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 OX NCBI\_TaxID=1358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Perreten V., Schwarz F., Boeglin M., Cresta L., Dasen G., Teuber M.;  
 RT "Antibiotic resistance spread in food.";  
 RL Nature 391:801-802(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21156959; PubMed=11257023;  
 RA Perreten V., Schwarz F., Teuber M., Levy S.B.;  
 RT "Mdt(A), a New Efflux Protein Conferring Multiple Antibiotic  
 Resistance in Lactococcus lactis and Escherichia coli.";  
 RL Antimicrob. Agents Chemother. 45:1109-1114(2001).  
 DR EMBL; X92946; CAA63504.1; -;  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 314 AA; 37078 MW; 8533511F7384C177 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 RDFPKYV 172  
 |||||  
 Db 182 RDFPKYV 188

RESULT 92

O970H0 PRELIMINARY; PRT; 315 AA.

AC O970H0;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein ST1624.  
 GN ST1624.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=111955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=JCM 10545 / 7;  
 RC MEDLINE=21456156; PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Kudoh R., Yamazaki J., Kushiya N., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yanagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermophilic

RT Crenarchaeon, Sulfolobus tokodaii strain7.",  
 RL DNA Res. 8:123-140(2001).  
 DR EMBL; AP000987; BAB66703.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 315 AA; 35692 MW; 3749E62D1671F43B CRC64;

Query Match 2.1%; Score 7; DB 17; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 KEPPTA 65  
 |||||  
 Db 33 KEPPTA 39

RESULT 93

O06996 PRELIMINARY; PRT; 320 AA.

AC O06996;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein yvdo.  
 GN YVDO.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Denizot F.C.;  
 RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=168;  
 RC MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Gim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengraber T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis.";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z94043; CAB08044.1; -;  
 DR EMBL; Z99121; CAB15458.1; -;

DR PIR; F70034; F70034.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
 DR InterPro; IPR002641; Patatin.  
 DR Pfam; PF01734; Patatin; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 320 AA; 35356 MW; 15909FEA37BC3125 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 ADLQLID 90  
 Db 103 ADLQLID 109  
 |||||

## RESULT 94

ID Q84BB3 PRELIMINARY; PRT; 321 AA.  
 AC Q84BB3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE HRPQ.  
 GN HRPQ.  
 OS Erwinia chrysanthemi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Pectobacterium.  
 OX NCBI\_TaxID=556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rojas C.M., Ham J.H., Kim J.F., Bear S.V., Collmer A.;  
 RT "Erwinia chrysanthemi hrp/hrc genes and flanking regions";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AV169276; AAO34611.1; -.  
 DR InterPro; IPR008984; SMAD\_FHA.  
 SQ SEQUENCE 321 AA; 35671 MW; 70FEAD778E7A7570 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 PPTA 67  
 Db 124 PPTA 130  
 |||||

## RESULT 95

ID O74237 PRELIMINARY; PRT; 322 AA.  
 AC O74237;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Xylose reductase (EC 1.1.1.21).  
 GN XLR.  
 OS Candida tenuis.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=45596;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hacker B., Habenicht A., Kies M., Mattes R.;  
 RT "Xylose Utilisation: Cloning and Characterization of the Xylose  
 Reductase from Candida tenuis";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF074484; AAC25601.1; -.  
 DR PDB; 1UEZ; 27-SEP-02.  
 DR PDB; 1K8C; 27-SEP-02.  
 DR GO; GO:0004032; F:aldehyde reductase activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR InterPro; IPR001395; Aldo/ket red.  
 DR Pfam; PF00248; aldo ket red; 1.  
 DR PRINTS; PR00069; ALDKETREDTASE.  
 DR Prodom; PD000288; Aldo/ket red; 1.  
 DR PROSITE; PS00798; ALDOKETO\_REDUCTASE\_1; 1.  
 DR PROSITE; PS00662; ALDOKETO\_REDUCTASE\_2; 1.  
 DR PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 322 AA; 36021 MW; 4C74A8FBC9357690 CRC64;

Query Match 2.1%; Score 7; DB 3; Length 322;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 QPKLIEF 304  
 Db 200 QPKLIEF 206  
 |||||

## RESULT 96

ID Q8LIF3 PRELIMINARY; PRT; 322 AA.  
 AC Q8LIF3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (P0503D09.26 protein).  
 GN OJ1316.A04.9 OR P0503D09.26.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC  
 clone-OJ1316.A04.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC  
 clone:P0503D09.";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003822; BAC06992.1; -.  
 DR EMBL; AP005455; BAC16736.1; -.  
 DR Gramene; O8LIF3; -.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR004892; Mo25.  
 DR Pfam; PF03204; Mo25; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 322 AA; 37091 MW; 99434DFA7C2DCD21 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 322;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 FDIASDA 183  
 Db 177 FDIASDA 183  
 |||||

## RESULT 97

ID Q97YG7 PRELIMINARY; PRT; 332 AA.  
 AC Q97YG7;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Endoglucanase precursor (EC 3.2.1.4).

GN SSO1354.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]\_TaxID=2287;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awevez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 DR EMBL; AB06749; AAK41590.1; -;  
 DR FIR; G90291; G90291.  
 DR GO; GO:0008810; F:cellulase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO; GO:000272; P:polysaccharide catabolism; IEA.  
 DR InterPro; IPR008985; Conn\_like\_1ec\_g1.  
 DR InterPro; IPR002594; Glyco\_hydro\_12.  
 DR Pfam; PF01670; Glyco\_hydro\_12; 1.  
 DR ProbDom; PD004316; Glyco\_hydro\_12; 1.  
 KW Hydrolase; Glycosidase; Complete proteome.  
 SQ SEQUENCE 332 AA; 37295 MW; F896165AE4926B9E CRC64;

Query Match 2.1%; Score 7; DB 17; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 FDIASDA 183

DB 184 FDIASDA 190

RESULT 98

P77472  
 ID P77472 PRELIMINARY; PRT; 334 AA.  
 AC P77472;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TNPA.  
 GN TNPA.  
 OS Escherichia coli.  
 OG Plasmid IncP-beta R751.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]\_TaxID=562;  
 RP SEQUENCE FROM N.A.  
 RA Thomas C.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Davidson P., Thomas C.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Davidson P., Thomas C.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Thomas C.M.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U67194; AAC64435.1; -;  
 DR EMBL; U60777; AAB03486.1; -;  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR003346; Transposase\_9.  
 DR InterPro; IPR002525; Transposase\_9.  
 DR Pfam; PF02371; Transposase\_20; 1.  
 DR Pfam; PF01548; Transposase\_9; 1.  
 KW Plasmid.

SQ SEQUENCE 334 AA; 37541 MW; 5935A9489C25EAD5 CRC64;  
 Query Match 2.1%; Score 7; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 PKLIEFL 305

DB 36 PKLIEFL 42

RESULT 99

P77600  
 ID P77600 PRELIMINARY; PRT; 334 AA.  
 AC P77600;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TNPA.  
 GN TNPA.  
 OS Escherichia coli.  
 OG Plasmid IncP-beta R751.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]\_TaxID=562;  
 RP SEQUENCE FROM N.A.  
 RA Davidson P., Thomas C.M.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Thomas C.M.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Davidson P., Thomas C.M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Thomas C.M.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U67194; AAC64438.1; -;  
 DR EMBL; U60777; AAB03489.1; -;  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR003346; Transposase\_9.  
 DR InterPro; IPR002525; Transposase\_9.  
 DR Pfam; PF02371; Transposase\_20; 1.  
 DR Pfam; PF01548; Transposase\_9; 1.  
 KW Plasmid.

SQ SEQUENCE 334 AA; 37449 MW; 464365D66CCF2E3 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 PKLIEFL 305

DB 36 PKLIEFL 42

RESULT 100

O8GGX8  
 ID O8GGX8 PRELIMINARY; PRT; 334 AA.  
 AC O8GGX8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Transposase.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M634;
RA Essa A.M., Julian D.J., Kidd S.P., Brown N.L., Hobman J.L.;
RT "Mercury resistance in Enterobacteriaceae from the pre-antibiotic
RT era.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457211; AAN87557.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR003346; Transposase_20.
DR InterPro; IPR002525; Transposase_9.
DR Pfam; PF02371; Transposase_20; 1.
DR Pfam; PF01548; Transposase_9; 1.
SQ SEQUENCE 334 AA; 37467 MW; FF745E317070903B CRC64;

Query Match 2.1%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 101
Q8GGX3 PRELIMINARY; PRT; 334 AA.
AC Q8GGX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transposase.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M634;
RA Essa A.M., Julian D.J., Kidd S.P., Brown N.L., Hobman J.L.;
RT "Mercury resistance in Enterobacteriaceae from the pre-antibiotic
RT era.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457211; AAN87557.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR003346; Transposase_20.
DR InterPro; IPR002525; Transposase_9.
DR Pfam; PF02371; Transposase_20; 1.
DR Pfam; PF01548; Transposase_9; 1.
SQ SEQUENCE 334 AA; 37423 MW; CC86FE7FB842BA0D CRC64;

Query Match 2.1%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 102
Q84FH4 PRELIMINARY; PRT; 334 AA.
AC Q84FH4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Transposase.
GN TNP.
OS Acinetobacter baumannii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=470;
RN [1]
RP SEQUENCE FROM N.A.
RA Ribera A., Roca I., Ruiz J., Gibert I., Vila J.;
RT "Partial characterization of a transposon containing the TET A
RT determinant in one clinical isolate of Acinetobacter baumannii.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY196695; AAO38184.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR003346; Transposase_20.
DR Pfam; PF02371; Transposase_20; 1.
SQ SEQUENCE 334 AA; 37467 MW; D629E2A1933B8F73 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 103
Q84D08 PRELIMINARY; PRT; 334 AA.
AC Q84D08;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TnpA.
GN TnpA.
OS Serratia marcescens.
OG Plasmid pDUI358.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RA Essa A.M.M., Hobman J.L., Brown N.L.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY225348; AAO85259.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR003346; Transposase_20.
DR Pfam; PF02371; Transposase_20; 1.
KW Plasmid.
SQ SEQUENCE 334 AA; 37467 MW; D629E2A1933B8F73 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305
Db 36 PKLIEFL 42

RESULT 104
Q84D05 PRELIMINARY; PRT; 334 AA.
AC Q84D05;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```

QY      299 PKLIEFL 305
DB      36 PKLIEFL 42
      |||||
      |||||

RESULT 106
Q7MUV2      PRELIMINARY;      PRT;      334 AA.
AC Q7MUV2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transposase.
OG Klebsiella pneumoniae.
OG Plasmid pRMH760.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RX [1]
RP SEQUENCE FROM N.A.
RA Partridge S.R., Hall R.M.;
RT "The IS1111 family members IS4321 and IS5075 have sub-terminal
RT inverted repeats and target the terminal inverted repeats of Tn21
RT family transposons.";
RL Submitted (FEB-2003) to the ENBL/GenBank/DBBJ databases.
DR ENBL; AY242532; AAO48503.1; -.
KW Plasmid.
SQ SEQUENCE      334 AA; 37525 MW;  D50466A1926FCB73 CRC64;

Query Match      2.1%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred.No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      299 PKLIEFL 305
DB      36 PKLIEFL 42
      |||||
      |||||

RESULT 107
Q93516      PRELIMINARY;      PRT;      334 AA.
AC Q93516;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase.
OG HCMV.209C.
OG Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608;
RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leach S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RP "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RL ENBL; AL513383; CAD09795.1; -.
DR GO; GO:0045821; C:extrachromosomal DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.

```



DR InterPro; IPR003346; Transposase\_20.  
 DR InterPro; IPR002525; Transposase\_9.  
 DR Pfam; PF02371; Transposase\_20; 1.  
 DR Pfam; PF01548; Transposase\_9; 1.  
 DR Plasmid; Complete proteome\_9.  
 SQ SEQUENCE 334 AA; 37467 MW; 0D3D1D57DD8D0603 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
 Db 36 PKLIEFL 42

RESULT 108  
 Q934H6 PRELIMINARY; PRT; 334 AA.  
 AC Q934H6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative transposase.  
 GN HCM1.236.  
 OS Salmonella typhi.  
 OG Plasmid pHCMI.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Baoham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 PT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 DR EMBL; AL513383; CAD09818.1; -.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004803; F:transposase activity; IEA.  
 DR GO; GO:0006313; P:DNA transposition; IEA.  
 DR InterPro; IPR003346; Transposase\_20.  
 DR InterPro; IPR002525; Transposase\_9.  
 DR Pfam; PF02371; Transposase\_20; 1.  
 DR Pfam; PF01548; Transposase\_9; 1.  
 DR Plasmid; Complete proteome\_9.  
 KW Plasmid; Complete proteome\_9.  
 SQ SEQUENCE 334 AA; 37467 MW; D629E2A1933B8F73 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
 Db 36 PKLIEFL 42

RESULT 109  
 Q97X08 PRELIMINARY; PRT; 334 AA.  
 ID Q97X08  
 AC Q97X08;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Endoglucanase precursor (EC 3.2.1.4).  
 GN SS01949.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 DR EMBL; AE006803; AAK42142.1; -.  
 DR PIR; G90360; G90360.  
 DR GO; GO:0008810; F:cellulase activity; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO; GO:0000272; P:polysaccharide catabolism; IEA.  
 DR InterPro; IPR008985; ConA like lec gl.  
 DR InterPro; IPR002594; Glyco\_hydro\_12.  
 DR Pfam; PF01670; Glyco\_hydro\_12; 1.  
 DR ProDom; PD004316; Glyco\_hydro\_12; 1.  
 DR Hydrolase; Glycosidase; Complete proteome.  
 SQ SEQUENCE 334 AA; 37619 MW; CB3F613936393A48 CRC64;

Query Match 2.1%; Score 7; DB 17; Length 334;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 FDIASDA 183  
 Db 186 FDIASDA 192

RESULT 110  
 Q98AA6 PRELIMINARY; PRT; 335 AA.  
 ID Q98AA6  
 AC Q98AA6;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ABC transporter, ATP-binding component.  
 GN MLR6085.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RL Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003008; BAB52430.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.

DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00005; ABC tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 335 AA; 35983 MW; 8CA1B1AB9649CA56 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLGEL 235  
 |||||  
 Db 199 LKLGEL 205

RESULT 111  
 Q831E0 PRELIMINARY; PRT; 335 AA.  
 AC Q831E0;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Conserved domain protein.  
 GN EF2571.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V583 / ATCC 700802;  
 RX MEDLINE=22550857; PubMed=12663927;  
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beaman M.,  
 RA Dagherty S., Desoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,  
 RA Vamathevan J., Tran B.T., Upton J., Hansen T., Shetty J., Khouri H.,  
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;  
 RT "Role of mobile DNA in the evolution of vancomycin-resistant  
 RT Enterococcus faecalis".  
 RL Science 299:2071-2074 (2003).  
 DR EMBL; AE016955; AA082282.1; -.  
 DR TIGR; EF2571; -.  
 DR InterPro; IPR003777; DUF182.  
 DR Pfam; PF02625; XdhC\_CoxI; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 335 AA; 37306 MW; 225BAGD143BDB792 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 ALLEKQD 32  
 |||||  
 Db 11 ALLEKQD 17

RESULT 112  
 Q7X1Q4 PRELIMINARY; PRT; 337 AA.  
 AC Q7X1Q4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein OJ1316\_A04.107.  
 GN OJ1316\_A04.107.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OX NCBI\_TaxID=39947;

DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00005; ABC tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 335 AA; 35983 MW; 8CA1B1AB9649CA56 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 FDIASDA 183  
 |||||  
 Db 177 FDIASDA 183

RESULT 113  
 Q85QAS PRELIMINARY; PRT; 339 AA.  
 AC Q85QAS;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Mitochondrial ribosomal protein VAR1.  
 GN VAR1.  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5478;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 2001;  
 RA Koszul R., Malpertuy A., Frangeul L., Bouchier C., Thierry A.,  
 RA Duthoy S., Winkler P., Ferris S., Hennequin C., Dujon B.;  
 RT "The complete mitochondrial genome sequence of the pathogenic yeast  
 RT Candida (Torulopsis) glabrata".  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 2001;  
 RA Frangeul L.;  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ511533; CA554417.1; -.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 DR InterPro; IPR007980; Yeast VAR1.  
 DR Pfam; PF05316; Yeast VAR1; 1.  
 KW Mitochondrion; Ribosomal protein.  
 SQ SEQUENCE 339 AA; 40828 MW; E56F22DB96DBEC06 CRC64;

Query Match 2.1%; Score 7; DB 8; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYIS 250  
 |||||  
 Db 180 IMTKYIS 186

RESULT 114  
 Q97L54 PRELIMINARY; PRT; 344 AA.  
 AC Q97L54;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative transcriptional regulator.  
 GN CAC0708.

OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]\_TaxID=1488;  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
RA Tatusov R.L., Sabathe P., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RA Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
bacterium Clostridium acetobutylicum";  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL; AE007586; AAK78685.1; -;  
DR PIR; B96987; B96987.  
DR GO; GO:0030246; F:carbohydrate binding; IEA.  
DR GO; GO:0030528; F:transcription regulator activity; IEA.  
DR InterPro; IPR007324; Sugar-bind.  
DR Pfam; PF04198; Sugar-bind; 1.  
KW Complete proteome.  
SQ SEQUENCE 344 AA; 38112 MW; 8898505500CE5D2D CRC64;

Query Match 2.1%; Score 7; DB 16; Length 344;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EIVKILK 22  
|||||  
Db 335 EIVKILK 341

RESULT 115  
Q9F0B0 PRELIMINARY; PRT; 345 AA.  
AC Q9F0B0;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE HrpZ.  
GN HRPZ.  
OS Pseudomonas syringae (pv. phaseolicola).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=319;  
RN [1]\_TaxID=319;  
RP SEQUENCE FROM N.A.  
RC STRAIN=1302A;  
RX MEDLINE=21065167; PubMed=11134504;  
RA Lee J., Klusener B., Tsiamis G., Stevens C., Neyt C., Tampakaki A.P.,  
RA Panopoulos N.J., Noller J., Weiler E.W., Cornelis G.R.,  
RA Mansfield J.W., Nurnberger T.;  
RT "HrpZeph from the plant pathogen Pseudomonas syringae pv.  
phaseolicola binds to lipid bilayers and forms an ion-conducting pore  
in vitro";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:289-294(2001).  
DR EMBL; AF268940; AAF99292.1; -;  
DR InterPro; IPR006961; HrpZ.  
DR Pfam; PF04877; HRPZ; 1.  
SQ SEQUENCE 345 AA; 35250 MW; A4B0B23A67268CA2 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 345;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AQLAQEL 73  
|||||  
Db 42 AQLAQEL 48

RESULT 116  
Q52481

ID Q52481 PRELIMINARY; PRT; 345 AA.  
AC Q52481;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE HrpZ protein.  
DE HRPZ.  
OS Pseudomonas syringae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=317;  
RN [1]\_TaxID=317;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Race 4;  
RX MEDLINE=96025089; PubMed=7579616;  
RA Preston G., Huang H.C., He S.Y., Collmer A.;  
RT "The HrpZ proteins of Pseudomonas syringae pvs. syringae, glycinea,  
RT and tomato are encoded by an operon containing Yersinia ysc homologs  
RL and elicit the hypersensitive response in tomato but not soybean.";  
RL Mol. Plant Microbe Interact. 8:717-732(1995).  
DR EMBL; L41862; AAB00136.1; -;  
DR InterPro; IPR006961; HrpZ.  
DR Pfam; PF04877; HrpZ; 1.  
SQ SEQUENCE 345 AA; 35290 MW; B73FB324B7F28DDC CRC64;

Query Match 2.1%; Score 7; DB 2; Length 345;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AQLAQEL 73  
|||||  
Db 42 AQLAQEL 48

RESULT 117  
Q8L9L9 PRELIMINARY; PRT; 345 AA.  
AC Q8L9L9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation";  
RL Genome Biol. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RT "Full-length cDNA from Arabidopsis thaliana";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY088359; AAM65898.1; -;  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR004892; Mo25.  
DR Pfam; PF03204; Mo25; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 345 AA; 39841 MW; 2C46A3D3DEBB47AA CRC64;

Query Match 2.1%; Score 7; DB 10; Length 345;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 TFXDLT 192  
|||||

Db 130 TFKDLT 196

## RESULT 118

Q7TN44 ID Q7TN44 PRELIMINARY; PRT; 353 AA.

AC Q7TN44; DT 01-OCT-2003 (TREMBlrel. 25, Created)  
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE MRG55 G protein-coupled receptor.

GN MRG55.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RX MEDLINE=22810130; PubMed=12909716;

RA Zylka M.J., Dong X., Southwell A.L., Anderson D.J.;

RT "Atypical expansion in mice of the sensory neuron-specific MrG G

RL protein-coupled receptor family.";

DR Proc. Natl. Acad. Sci. U.S.A. 100:10043-10048 (2003).

DE EMBL; AF518243; AAQ08315.1; -.

KW Receptor.

SQ SEQUENCE 353 AA; 40125 MW; CE3685C2D3A6A9E0 CRC64;

## Query Match

Best Local Similarity 2.1%; Score 7; DB 11; Length 353;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 RQSLKL 232

Db 319 RQSLKL 325

## RESULT 119

Q9YET3 ID Q9YET3 PRELIMINARY; PRT; 357 AA.

AC Q9YET3; DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE 357AA long hypothetical CARBAMOYLPHOSPHATE synthetase.

GN APE0498.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;

OC Desulfurococaceae; Aeropyrum.

OX NCBI\_TaxID=56636;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K1;

RX MEDLINE=99110339; PubMed=10382966;

RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,

RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,

RT "Complete genome sequence of an aerobic hyper-thermophilic

Crenarchaeon, Aeropyrum pernix K1.";

RL DNA Res. 6:83-101 (1999).

DR EMBL; AF000059; BAA79463.1; -.

DR PIR; C72746; C72746.

DR InterPro; IPR007356; DUF425.

DR Pfam; PF04243; DUF425; 1.

KW Complete proteome.

SQ SEQUENCE 357 AA; 40622 MW; FA21C414838695F4 CRC64;

## Query Match

Best Local Similarity 2.1%; Score 7; DB 17; Length 357;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILLK 296  
Db 269 IVEILLK 275

## RESULT 120

Q8FVE4 ID Q8FVE4 PRELIMINARY; PRT; 358 AA.

AC Q8FVE4; DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Ornithine cyclodeaminase.

GN ARCB OR BR0899.

OS Brucella suis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI\_TaxID=29461;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;

RX MEDLINE=22247741; PubMed=12271122;

RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,

RA Dougherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,

RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,

RA Hoover D.B., Lindler L.E., Hailing S.M., Boyle S.M., Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between

animal and plant pathogens and symbionts.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).

DR EMBL; AE014583; AAN34071.1; -.

DR TIGR; BR0899; -.

DR InterPro; IPR003462; ODC\_Mu\_crystall.

DR Pfam; PF02423; ODC\_Mu\_crystall; 1.

KW Complete proteome.

SQ SEQUENCE 358 AA; 39299 MW; DA2F02B1171B4C81 CRC64;

## Query Match

Best Local Similarity 2.1%; Score 7; DB 16; Length 358;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EKLQSE 220

Db 346 EKLQSE 352

## RESULT 121

Q8RMW1 ID Q8RMW1 PRELIMINARY; PRT; 359 AA.

AC Q8RMW1; DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE FixB.

GN FIXB.

OS Azospirillum brasilense.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;

OC Rhodospirillaceae; Azospirillum.

OX NCBI\_TaxID=192;

RN [1]

RP SEQUENCE FROM N.A.

RC Gross J., Sperotto R.A., Vedoy C., Passaglia L.M.P., Schrank I.S.;

RT "The Azospirillum brasilense fixABCX operon: nucleotide sequence.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF489443; AAM00919.1; -.

DR GO; GO:0005489; F:electron transporter activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR013308; ETF\_alpha.

DR Pfam; PF00766; ETF\_alpha; 1.

DR PROSITE; PS00696; ETF\_ALPHA; 1.

DR PROSITE; PS00696; ETF\_ALPHA; 1.

SQ SEQUENCE 359 AA; 38708 MW; 89EBCCE6498F3D54 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPNLKL 257  
 DB 245 KPNLKL 251

RESULT 122

QY 017497 PRELIMINARY; PRT; 359 AA.

AC 017497;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Aldolase.  
 OS Branchiostoma belcheri (Amphioxus).  
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
 CC Branchiostoma.  
 CX NCBI\_TaxID=7741;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Kuba M., Yatsuki H., Kusakabe T., Takasaki Y., Nikoh N., Miyata T.,  
 RA Yanaguchi T., Hori K.;  
 RT "Molecular Evolution of Amphioxus Fructose-1,6-bisphosphate  
 RT Aldolase";  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB005035; BAA21101.1; -;  
 DR HSSP; P00883; 1ADO.  
 DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.  
 DR GO; GO:0006036; P:glycolysis; IEA.  
 DR InterPro; IPR000741; Aldolase I.  
 DR Pfam; PF00274; glycolytic enzy; 1.  
 DR ProDom; PD01128; Aldolase I; 1.  
 DR PROSITE; PS00158; ALDOLASE\_CLASS\_I; 1.  
 SQ SEQUENCE 359 AA; 38493 MW; 9A0374AE31EBE18 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKDN 24  
 DB 91 VKILKDN 97

RESULT 123

QY 08VCY1 PRELIMINARY; PRT; 359 AA.

AC 08VCY1;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ornithine cyclodeaminase (EC 4.3.1.12).  
 GN BMEI0397.  
 OS Brucella melitensis.  
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Brucellaceae; Brucella.  
 CX NCBI\_TaxID=29459;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / Biotype 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muijer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Havelkorn R., Kyrpides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).

DR EMBL; AE009677; AAL53639.1; -;  
 DR PIR; AD3559; AD3559.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0008473; F:ornithine cyclodeaminase activity; IEA.  
 DR InterPro; IPR003462; ODC\_Mu\_crystall.  
 DR Pfam; PF02423; ODC\_Mu\_crystall; 1.  
 KW Lyase; Complete proteome.  
 SQ SEQUENCE 359 AA; 39460 MW; F29E90021EF950B5 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EKLQSE 220  
 DB 347 EKLQSE 353

RESULT 124

QY 09KS99 PRELIMINARY; PRT; 365 AA.

AC 09KS99;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Amino acid ABC transporter, permease protein.  
 GN VCI360.  
 OS Vibrio cholerae.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 CC Vibrionaceae; Vibrio.  
 CX NCBI\_TaxID=666;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae";  
 RL Nature 406:477-483(2000).  
 CC -1- FUNCTION: PROBABLY PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT  
 CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF THE  
 CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-  
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.

DR EMBL; AE004215; AAF94518.1; -;  
 DR PIR; F82210; F82210.  
 DR TIGR; VCI360; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000515; BPD\_transp.  
 DR InterPro; IPR00528; BPD\_transp; 1.  
 DR Pfam; PF00528; BPD\_transp; 1.  
 DR PROSITE; PS00402; BPD\_TRANSP\_INN\_MEMBR; 1.  
 KW Transmembrane; Transport; Complete proteome.  
 SQ SEQUENCE 365 AA; 40698 MW; C783C9BAA49D24FD CRC64;

Query Match 2.1%; Score 7; DB 16; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLIVTILI 83  
 DB 157 GLIVTILI 163

RESULT 125

```

Q8D999
ID Q8D999 PRELIMINARY; PRT; 365 AA.
AC Q8D999;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC-type amino acid transport system, permease component.
GN V12705.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016806; AA011051.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; 1.
RW Complete proteome.
SQ SEQUENCE 365 AA; 41096 MW; DE355603114A41A8 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 GLLVTLI 83
Db 157 GLLVTLI 163

RESULT 126
Q8VZN7 PRELIMINARY; PRT; 368 AA.
AC Q8VZN7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative peptide transporter.
GN At1G64500.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Yamada K., Banno F., Chang E., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At1g64500 (GI:15217659).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY063982; AAJ36338.1; -
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002109; Glutaredoxin.
DR Pfam; PF00462; Glutaredoxin; 1.
RW Hypothetical protein.
SQ SEQUENCE 368 AA; 40986 MW; 20A65B0127DA24A4 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 368;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 EVVSKSL 46
Db 76 EVVSKSL 82

RESULT 127
Q9SGW5 PRELIMINARY; PRT; 368 AA.
AC Q9SGW5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FIN19.7 (Hypothetical protein) (Putative peptide transporter
DE protein).
GN FIN19.7 OR At1G64500.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Tambunga G., Altafi H., Bei Q., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thavari A., Toriumi M., Vayaberg M., Yu G., Federspiel N.A.,
RA Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC FIN19 from chromosome
RT I."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Yamada K., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene FIN19.7 (GI:6633811).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009519; AAF19670.1; -
DR EMBL; AY056260; AAL07109.1; -
DR EMBL; BT002750; AAO22579.1; -
DR FR; G96668; G96668.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002109; Glutaredoxin.
DR Pfam; PF00462; Glutaredoxin; 1.
RW Hypothetical protein.
SQ SEQUENCE 368 AA; 41016 MW; 39F8B0127C937A4 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 368;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 EVVSKSL 46
Db 76 EVVSKSL 82

```

```
RESULT 128
Q87P96
ID Q87P96 PRELIMINARY; PRT; 368 AA.
AC Q87P96;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amino acid ABC transporter, permease protein.
GN VP1622.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=RMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Tajima Y., Nishida M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005078; BAC59885.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp. 1.
DR PROSITE; PS00402; BPD_TRANS_INN_MEMBER; 1.
KW Complete proteome.
SQ SEQUENCE 368 AA; 41399 MW; A09AE5073B410381 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 368;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTLI 83
DB 157 GLLVTLI 163

RESULT 129
Q9RDY8
ID Q9RDY8 PRELIMINARY; PRT; 369 AA.
AC Q9RDY8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
DE Hypothetical protein.
OS Legionella pneumophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=446;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=RC1;
RX MEDLINE=20496696; PubMed=11043980;
RA Lueneberg E., Zetzmann N., Hartmann M., Knirel Y.A., Koolstra O.,
RA Zaehring U., Helbig J., Frosch M.;
RT "Cloning and functional characterization of a 30 kb gene locus
RT required for lipopolysaccharide biosynthesis in Legionella
RT pneumophila.";
RL Int. J. Med. Microbiol. 290:37-49(2000).
DR EMBL; AJ007311; CAB65199.1; -
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 42303 MW; 40955E0A9DFA2624 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 EIVKILK 22
DB 18 EIVKILK 24

RESULT 130
Q7YQJ9
ID Q7YQJ9 PRELIMINARY; PRT; 369 AA.
AC Q7YQJ9;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ubiquitin-conjugating enzyme UBCi.
DE Ubiquitin-conjugating enzyme UBCi.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Implantation stage endometrium;
RA Melner M.H., Ducharme N.A., Brash A.R., Winfrey V.P., Olson G.E.;
RT "Differential Expression of Genes in the Endometrium at Implantation:
RT Upregulation of a Novel Member of the E2 Class of Ubiquitin
RT Conjugating Enzymes.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY303051; AAP93920.1; -
SQ SEQUENCE 369 AA; 42191 MW; 3B88532499960CF3 CRC64;

Query Match 2.1%; Score 7; DB 6; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILK 30
DB 173 NLAILK 179

RESULT 131
Q8WVN8
ID Q8WVN8 PRELIMINARY; PRT; 375 AA.
AC Q8WVN8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to NICE-5 protein (EC 6.3.2.19) (Ubiquitin-conjugating enzyme
DE E2) (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; BC017708; AAHI7708.1; -
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR006575; RWD.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UBQ_con; 1.
DR Pfam; PD000461; UBQ_conjugat; 1.
```



```

DR SMART; SMO0591; RWD; 1.
DR SMART; SMO0212; UBCC; 1.
DR PROSITE; PS0127; UBIQUITIN CONJUGAT 2; 1.
KW Ligase; Ub1 conjugation pathway.
SQ SEQUENCE 375 AA; 42818 MW; 7DE07315E89178A3 CRC64;

Query Match      2.1%; Score 7; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 NLAILEK 30
Db      179 NLAILEK 185

RESULT 132
Q8LAH2 PRELIMINARY; PRT; 377 AA.
ID Q8LAH2
AC Q8LAH2
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Polyalacturonase-like protein.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0 (2002).
[2]
RN SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGLACTURONASES).
CC EMBL; AY087812; AM65366.1;
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016798; F:hydrolase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PBH1.
DR InterPro; IPR000508; Peptidase_S26.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SMO0710; PBH1; 4.
DR PROSITE; PS00761; SPASE I 3; 1.
KW Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 377 AA; 40865 MW; 4DD25136A4642102 CRC64;

Query Match      2.1%; Score 7; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      301 LIEFLSS 307
Db      112 LIEFLSS 118

RESULT 133
Q9WY73 PRELIMINARY; PRT; 377 AA.
ID Q9WY73
AC Q9WY73
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F3E22.9 protein.
GN F3E22.9.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F3E22 genomic sequence."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGLACTURONASES).
CC EMBL; AC023912; AAF63821.1;
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016798; F:hydrolase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PBH1.
DR InterPro; IPR000508; Peptidase_S26.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SMO0710; PBH1; 4.
DR PROSITE; PS00761; SPASE I 3; 1.
KW Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 377 AA; 40807 MW; B12FA24567860E1 CRC64;

Query Match      2.1%; Score 7; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      301 LIEFLSS 307
Db      112 LIEFLSS 118

RESULT 134
Q8K2Z8 PRELIMINARY; PRT; 378 AA.
ID Q8K2Z8
AC Q8K2Z8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to NICE-5 protein (EC 6.3.2.19) (Ubiquitin-conjugating enzyme
DE E2) (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -1- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; BC029111; AAH29111.1;
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

```

DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS0127; UBIQUITIN CONJUGAT\_2; 1.  
 KW Ligase; Ubl conjugation pathway.  
 SQ SEQUENCE 378 AA; 42923 MW; D1690A9C4BC6DBDC CRC64;

Query Match 2.1%; Score 7; DB 11; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
 |||||  
 DB 182 NLAILEK 188

## RESULT 135

Q8BUN2 PRELIMINARY; PRT; 378 AA.  
 AC Q8BUN2;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to NICE-5 protein homolog.  
 GN 3010021M2IRIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Hippocampus;  
 RX MEDLINE=22354683; PubMed=12468851;

RA The FANTOM Consortium,

RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

DR EMBL; AK083216; BAC38813.1; -

DR MGI; 1924088; 3010021M2IRIK.

DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR000608; UBQ\_conjugat.

DR Pfam; PF00179; UQ\_con; 1.

DR ProDom; PD000461; UBQ\_conjugat; 1.

DR SMART; SM00212; UBCC; 1.

DR PROSITE; PS0127; UBIQUITIN CONJUGAT\_2; 1.

SQ SEQUENCE 378 AA; 42950 MW; 4A81CA85400A1313 CRC64;

## Query Match

Best Local Similarity 2.1%; Score 7; DB 11; Length 378;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
 |||||  
 DB 182 NLAILEK 188

## RESULT 136

Q19391 PRELIMINARY; PRT; 379 AA.  
 AC Q19391;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein F13D11.4.  
 GN F13D11.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RA "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018 (1998).

[2]

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Fulton L.;  
 RA "The sequence of C. elegans cosmid F13D11.";  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

[3]

RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U40939; AA81703.2; -.  
 DR F1R; T16059; T16059.  
 DR WormPep; F13D11.4; CE30950.  
 KW Hypothetical protein.  
 SQ SEQUENCE 379 AA; 42494 MW; 221F2D4EB8B3138 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 VEILLKN 297  
 |||||  
 DB 58 VEILLKN 64

## RESULT 137

Q8E1Y3 PRELIMINARY; PRT; 381 AA.  
 AC Q8E1Y3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Galactokinase.  
 GN GALK OR SO0694.  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 OX NCBI\_TaxID=70863;

[1]

RN SEQUENCE FROM N.A.

RC STRAIN=MR-1.

RX MEDLINE=22297686; PubMed=12368813;

RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
 RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,  
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,  
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 Shewanella oneidensis.";

RL Nat. Biotechnol. 20:1118-1123 (2002).

DR EMBL; AE015515; AA53772.1; -.

DR TIGR; SO694; -.

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004335; F:galactokinase activity; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0006012; P:galactose metabolism; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR GO; GO:0016310; P:phosphorylation; IEA.

DR InterPro; IPR00705; Galactokinase.

DR InterPro; IPR001174; Galkinase.

DR InterPro; IPR006203; GMPKase ATP.  
 DR InterPro; IPR006204; GMP kinase.  
 DR InterPro; IPR006206; Mxv\_galkinase.  
 DR Pfam; PF00288; GMP\_kinases; 1.  
 DR PRINTS; PR00473; GALTOKINASE.  
 DR PRINTS; PR00960; LMBPPTOKINASE.  
 DR PRINTS; PR00959; MEVGALKINASE.  
 DR TIGRFAMs; TIGR00131; gal\_kin; 1.  
 DR PROSITE; PS00627; GMP\_KINASES\_ATP; 1.  
 KW Kinase; Complete proteome.  
 SQ SEQUENCE 381 AA; 41510 MW; C799B46FC8B544E2 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 AVAQLAQ 71  
 |||||  
 Db 154 AVAQLAQ 160

RESULT 138  
 Q9BKP3 PRELIMINARY; PRT; 386 AA.  
 AC Q9BKP3;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN H19W22.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.,";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.; Wamsley P.;  
 RT "The sequence of C. elegans cosmid H19W22.,";  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RT "Direct Submission,";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF040648; AAK26140.1; -.  
 DR WormPep; H19W22.2c; CE26978.  
 KW Hypothetical protein.  
 SQ SEQUENCE 386 AA; 43533 MW; BBE7BC17601E57D CRC64;

Query Match 2.1%; Score 7; DB 5; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AQLAQEL 73  
 |||||  
 Db 274 AQLAQEL 280

RESULT 139  
 Q8EY27 PRELIMINARY; PRT; 387 AA.  
 ID Q8EY27;  
 AC Q8EY27;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Probable intercellular adhesion protein C.  
 GN LB030.  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB011592; AAN51589.1; -.  
 DR GO; GO:0016747; P:transferase activity, transferring groups o. . .; IEA.  
 DR InterPro; IPR002656; Acyl\_transf\_3.  
 DR Pfam; PF01757; Acyl\_transf\_3; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 387 AA; 44605 MW; 4B3A7BC91D8661FB CRC64;

Query Match 2.1%; Score 7; DB 16; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
 |||||  
 Db 138 PKLIEFL 144

RESULT 140  
 Q9JRN4 PRELIMINARY; PRT; 390 AA.  
 ID Q9JRN4;  
 AC Q9JRN4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative glycosyltransferase.  
 OS Actinobacillus actinomycetemcomitans (Haemophilus  
 actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Actinobacillus.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SUNYab 75;  
 RA Suzuki N.; Nakano Y.; Yoshida Y.; Nako H.; Yamashita Y.; Koga T.;  
 RT "Genetic analysis of the gene cluster for the synthesis of serotype a-  
 RT specific polysaccharide antigen in A. actinomycetemcomitans.,";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB046360; BAB03209.1; -.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0009058; P:biogenesis; IEA.  
 DR InterPro; IPR001296; Glyco\_transf\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Transferase.  
 SQ SEQUENCE 390 AA; 44354 MW; 2B9267E002D35382 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 IMTKYIS 250  
 |||||  
 Db 275 IMTKYIS 281

RESULT 141  
 Q8CX72 PRELIMINARY; PRT; 396 AA.  
 ID Q8CX72;  
 AC Q8CX72;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Antibiotic resistance protein (Antibiotic efflux protein).

GN OB3076.  
 OS Oceanobacillus iheyensis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
 OX NCBI\_TaxID=182710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
 RX MEDLINE=2220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama I.;  
 RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments";  
 RL Nucleic Acids Res. 30:3927-3935 (2002).  
 DR EMBL; AP004603; BAC15032.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR007114; MFS.  
 DR InterPro; IPR005828; Sub\_transporter.  
 DR Pfam; PF00083; sugar\_tr; 1.  
 DR PROSITE; PS50850; MFS; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 396 AA; 42992 MW; 60BE167BC1C51569 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 283 SPHKTQP 289  
 Db 201 SPHKTQP 207

RESULT 142  
 Q9FL16 PRELIMINARY; PRT; 399 AA.  
 AC Q9FL16;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Adenylate cyclase.  
 GN CYAL.  
 OS Thermosynechococcus elongatus.  
 OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.  
 OX NCBI\_TaxID=146786;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Katoh H., Nakahara Y., Ikeuchi M.;  
 RT "Cloning and expression of the cyaI gene from thermophilic  
 RT Thermosynechococcus (formerly Synchococcus) elongatus BP-1";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB052847; BAB20624.1; -.  
 DR GO; GO:0004383; F:guanylate cyclase activity; IEA.  
 DR GO; GO:0016829; F:lyase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro; IPR00253; FHA.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR InterPro; IPR008984; SMAD\_FHA.  
 DR Pfam; PF00498; FHA; 1.  
 DR PROSITE; PS50006; FHA\_DOMAIN; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 KW Lyase.  
 SQ SEQUENCE 399 AA; 44620 MW; CF4957D52C3D7D1A CRC64;

Query Match 2.1%; Score 7; DB 2; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 130 LKGVEAP 136  
 Db 367 LKGVEAP 373

RESULT 143  
 Q9DGB0 PRELIMINARY; PRT; 399 AA.  
 AC Q9DGB0;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Adenylate cyclase.  
 GN TLH2410.  
 OS Synchococcus elongatus (Thermosynechococcus elongatus).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.  
 OX NCBI\_TaxID=32046;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BP-1;  
 RX MEDLINE=2225144; PubMed=12240834;  
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,  
 RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the thermophilic cyanobacterium  
 RT Thermosynechococcus elongatus BP-1";  
 RL DNA Res. 9:123-130 (2002).  
 DR EMBL; AP005377; BAC09962.1; -.  
 DR GO; GO:0004383; F:guanylate cyclase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro; IPR00253; FHA.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00498; FHA; 1.  
 DR SMART; SM00044; CYC; 1.  
 DR SMART; SM00240; FHA; 1.  
 DR PROSITE; PS50006; FHA\_DOMAIN; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 399 AA; 44620 MW; CF4957D52C3D7D1A CRC64;

Query Match 2.1%; Score 7; DB 16; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 LKGVEAP 136  
 Db 367 LKGVEAP 373

RESULT 144  
 Q92VE7 PRELIMINARY; PRT; 400 AA.  
 AC Q92VE7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein RB0759.  
 GN R30759 OR SNE21255.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OG Plasmid pSymb (megaplasmid 2).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396508; PubMed=11481431;  
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P., Gouzy J.,  
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,  
 RA Golding B., Puhler A.;  
 RT "The complete sequence of the 1,693-kb pSymb megaplasmid from the N2-  
 RT fixing endosymbiont Sinorhizobium meliloti";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).  
 DR EMBL; AL603644; CAC49159.1; -.

DR PIR; G95936; G95936.

DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 KW Plasmid; Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 400 AA; 42766 MW; 3F775E554D0F5F1F CRC64;

Query Match 2.1%; Score 7; DB 16; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 KASEEVS 43  
 Db 140 KASEEVS 146  
 |||||

RESULT 145

Q8BLI3 Q8BLI3 PRELIMINARY; PRT; 402 AA.

AC Q8BLI3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Dynamin-like 120 kDa protein (Fragment).  
 GN OPAL.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Retina;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,

the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).  
 DR EMBL; AK044657; BAC32021.1; -

DR MGD; MGI:192139; Opal.  
 DR GO; GO:0005739; C:mitochondrion; IDA.

DR InterPro; IPR001401; Dynamin.  
 DR Pfam; PF00350; dynamin; 1.  
 DR PRINTS; PR00195; DYNAMIN.

FT NON TER 402  
 SQ SEQUENCE 402 AA; 46232 MW; E660EC93BC32E74E CRC64;

Query Match 2.1%; Score 7; DB 11; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELIL 237  
 Db 210 LLGELIL 216  
 |||||

RESULT 146

Q7UWG3 Q7UWG3 PRELIMINARY; PRT; 402 AA.

AC Q7UWG3;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.

GN R2055  
 OS Rhodospirillum baltica.

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetacia; Pirellula.

OX NCBI\_TaxID=117;  
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O.; Kube M.; Bauer M.; Teeling H.; Lombardot T.

RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,

RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294136; CAD72400.1; -

KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 402 AA; 43264 MW; 4165BF55C907C188 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 LLKNQPK 300  
 Db 28 LLKNQPK 34  
 |||||

RESULT 147

Q9XAY2 Q9XAY2 PRELIMINARY; PRT; 407 AA.

AC Q9XAY2;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (Fragment).

OS Prevotella albensis.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Prevotellaceae; Prevotella.

OX NCBI\_TaxID=77768;  
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=M384;  
 RA McEwan N.R., Walker N.D., Wallace R.J.;

RT "A long open reading frame in the rumen bacterium Prevotella albensis  
 M384 has homology to unidentified open reading frames in other

organisms.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ243361; CAB46012.1; -

DR InterPro; IPR004813; Tetrpept\_transpt.  
 DR Pfam; PF03169; OPT; 1.

KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 407 AA; 42560 MW; 6FC6C897060762D9 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 407;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84  
 Db 222 LLVTLIA 228  
 |||||

RESULT 148

Q8M431 Q8M431 PRELIMINARY; PRT; 407 AA.

AC Q8M431;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NADH dehydrogenase subunit 5 (Fragment).

GN ND5.  
 OS Panthera tigris (Tiger).

OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Panthera.

OX NCBI\_TaxID=9694;  
 RN [1]

SEQUENCE FROM N.A.

RA Dubach J., Patterson B.D., Briggs M.B., Venzke K., Flammand J.,  
 RA Stander P., Scheepers L., Kays R.;

RT "Molecular genetic variation across the southern geographic range of  
 the African lion, Panthera leo.";

```
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385623; AAM46207.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR003916; NADH_oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
DR PRINTS; PR01434; NADHGHNGASE5.
DR NAD; Oxidoreductase, Ubiquinone, Mitochondrion.
KW NON_TER 407 407
FT NON_TER 407 407
SQ SEQUENCE 407 AA; 45157 MW; 29C65A27F1854410 CRC64;

Query Match 2.1%; Score 7; DB 8; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
Db 222 LLVTLIA 228
|||||

RESULT 149
Q8M432 PRELIMINARY; PRT; 407 AA.
AC Q8M432;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 5 (Fragment).
GN ND5.
OS Panthera leo (Lion).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Panthera.
OX NCBI_TaxID=9689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TSV3550;
RA Dubach J., Patterson B.D., Briggs M.B., Venzke K., Flammand J.,
RA Stander P., Scheepers L., Kays R.;
RT "Molecular genetic variation across the southern geographic range of
RT the African lion, Panthera leo."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385621; AAM46205.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR003916; NADH_oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR005806; Rieske_dom.
DR Pfam; PF00361; oxidored_q1; 1.
DR PRINTS; PR01434; NADHGHNGASE5.
DR NAD; Oxidoreductase, Ubiquinone, Mitochondrion.
KW NON_TER 1 1
FT NON_TER 407 407
SQ SEQUENCE 407 AA; 45183 MW; 70048488F626B884 CRC64;

Query Match 2.1%; Score 7; DB 8; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
Db 222 LLVTLIA 228
|||||

RESULT 150
Q7Z7E8 PRELIMINARY; PRT; 422 AA.
AC Q7Z7E8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-conjugating enzyme E2Q.
GN UBE2Q.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Altman M.E., Schulze E., Adham I., Koehler M., Engel W.;
RT "Isolation and characterization of the human UBE2Q gene and its murine
RT ortholog."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY112698; AAM60814.1; -.
SQ SEQUENCE 422 AA; 46127 MW; 106FF7B59DF65555 CRC64;

Query Match 2.1%; Score 7; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLALIEK 30
Db 226 NLALIEK 232
|||||

RESULT 151
Q7TSS2 PRELIMINARY; PRT; 422 AA.
AC Q7TSS2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ubiquitin-conjugating enzyme E2Q.
GN UBE2Q.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L29/SV;
RA Altman M.E., Schulze E., Adham I., Koehler M., Engel W.;
RT "Isolation and characterization of the human UBE2Q gene and its murine
RT ortholog."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY112699; AAM60815.1; -.
SQ SEQUENCE 422 AA; 46131 MW; FAA53C0CA87B8A5B CRC64;

Query Match 2.1%; Score 7; DB 11; Length 422;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLALIEK 30
Db 226 NLALIEK 232
|||||

RESULT 152
Q9ATH0 PRELIMINARY; PRT; 426 AA.
AC Q9ATH0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aminolevulinat dehydratase (EC 4.2.1.24) (Delta-aminolevulinic acid
DE dehydratase) (Porphobilinogen synthase) (ALADH).
GN ALAD.
OS Raphanus sativus (Radish).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Raphanus.
OX NCBI_TaxID=3726;
RN [1]
RP SEQUENCE FROM N.A.
RA Berthe T., Klein-Eude D., Balange A.P.;
RT "Study of 5-aminolevulinic dehydratase in radish seedlings: Are there
RT housekeeping and light-induced enzymes?";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 2 5-AMINOLEVULINATE = PORPHOBILINOGEN + 2
CC H(2)O.
CC -1- SUBUNIT: HOMOOCTAMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ALADH FAMILY.
CC EMBL; AF332195; AAK15323.1; -.
DR HSSP; P15002; 1B4E.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004655; F:porphobilinogen synthase activity; IEA.
DR GO; GO:0006783; F:heme biosynthesis; IEA.
DR InterPro; IPR001731; ALAD_dehydratase.
DR Pfam; PF00490; ALAD; 1.
DR PRINTS; PR00144; DALDHVDRTASE.
DR ProDom; PD002304; Alad dehydratase; 1.
DR PROSITE; PS00169; D_ALA_DEHYDRATASE; 1.
KW Lyase, Porphyrin biosynthesis.
SQ SEQUENCE 426 AA; 46552 MW; 0C239C00B77B1CAE CRC64;

Query Match 2.1%; Score 7; DB 10; Length 426;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LRRDKSP 267
DB 358 LRRDKSP 364
|||||
[1]

RESULT 153
Q97J34 PRELIMINARY; PRT; 436 AA.
ID Q97J34;
AC Q97J34;
RX MEDLINE=21359325; PubMed=11466286;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fe-S oxidoreductases.
GN CAC1286.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatunov R.L., Sabatthe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007641; AAK79257.1; -.
DR PIR; F97058; F97058.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR006467; MAb-like C.
DR InterPro; IPR007197; Radical SAM.
DR InterPro; IPR005839; UPF0004.
DR Pfam; PF04055; Radical SAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRfams; TIGR01579; MAb-like-C; 1.
DR TIGRfams; TIGR00089; TIGR00089; 1.
DR PROSITE; PS01278; UPF0004; 1.
KW Complete proteome.
SQ SEQUENCE 436 AA; 50205 MW; A3EF05AC290C93BF CRC64;

Query Match 2.1%; Score 7; DB 16; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILKD 23
DB 262 IVKILKD 268
|||||
[1]

RESULT 155
Q7VIN8 PRELIMINARY; PRT; 442 AA.
ID Q7VIN8;
AC Q7VIN8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FKBP-type peptidyl-prolyl cis-trans isomerase (Trigger factor).
GN TIG OR HH0566.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
```



```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droge M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003);
DR EMBL; AE017145; AAP77163.1; -.
KW Isomerase; Complete proteome.
SQ SEQUENCE 442 AA; 50710 MW; DB9377336D28834C CRC64;

Query Match 2.1%; Score 7; DB 16; Length 442;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 VEILLKN 297
Db 296 VEILLKN 302
|||||

RESULT 156
Q8T2G0 PRELIMINARY; PRT; 443 AA.
AC Q8T2G0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115608; AAL92378.2; -.
KW Hypothetical protein.
SQ SEQUENCE 443 AA; 51998 MW; 2BBD1D59F237B074 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 443;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIBFLSS 307
Db 8 LIBFLSS 14
|||||

RESULT 157
Q56568 PRELIMINARY; PRT; 444 AA.
AC Q56568;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VIRC.
GN VIRC.
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.

```

```

OX NCBI_TaxID=55601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NB10;
RX MEDLINE=96060845; PubMed=7590330;
RA Milton D.L., Norqvist A., Wolf-Watz H.;
RT "Sequence of a novel virulence-mediating gene, virC, from Vibrio
RT anguillarum.";
RL Gene 164:95-100(1995).
DR EMBL; U17054; AAA86985.1; -.
DR PIR; JC4348; JC4348.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
SQ SEQUENCE 444 AA; 51461 MW; 78F0B5901907717F CRC64;

Query Match 2.1%; Score 7; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 AQELYS 76
Db 130 AQELYS 136
|||||

RESULT 158
Q7SXY7 PRELIMINARY; PRT; 444 AA.
AC Q7SXY7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max J.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055195; AAH55195.1; -.
KW Hypothetical protein.
SQ SEQUENCE 444 AA; 51111 MW; 77898CC82F41BABE CRC64;

Query Match 2.1%; Score 7; DB 13; Length 444;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

```

ID Q7XUF9 PRELIMINARY; PRT; 453 AA.
AC Q7XUF9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE QJ991113.30.7 protein.
GN QJ991113.30.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Xu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.O., Guan J.P., Hong G.F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL62946; CAB41324.1; -.
SQ SEQUENCE 453 AA; 50532 MW; 6A8ABFD7D795AAA3 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 VLVADFL 202
DB 31 VLVADFL 37
|||||

```

```

RESULT 163
Q875W3 PRELIMINARY; PRT; 459 AA.
ID Q875W3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PH087 (Fragment).
OS Saccharomyces castellii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=27288;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CBS4309;
RX MEDLINE=22482865; PubMed=12594514;
RA Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;
RT "Yeast genome duplication was followed by asynchronous differentiation
of duplicated genes."
RL Nature 421:848-852(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CBS4309;
RA Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144952; AAC03515.1; -.
FT NON_TER 1
SQ SEQUENCE 459 AA; 48585 MW; 3EC9C765C08E9A23 CRC64;

Query Match 2.1%; Score 7; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLVT 81
DB 311 SSGLLVT 317
|||||

```

```

RESULT 164
Q8XLZ4 PRELIMINARY; PRT; 475 AA.
ID Q8XLZ4;
AC Q8XLZ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Two-component sensor histidine kinase.
GN CPE0896.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003188; BAB80602.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02518; HATPase_C; 1.
DR SMART; SM00387; HATPase_C; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 475 AA; 53567 MW; B889927345727429 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 475;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVSKSLQ 47
DB 22 EVSKSLQ 28
|||||

```

```

RESULT 165
Q9AWX3 PRELIMINARY; PRT; 476 AA.
ID Q9AWX3;
AC Q9AWX3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ribosomal RNA apurinic site specific lyase.
GN P0013G02.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Cryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0013G02."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002908; BAB32936.1; -.
DR Gramene; O9AWX3; -.
DR GO; GO:0016823; F:lyase activity; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR006566; FBD.
DR InterPro; IPR008945; Skp1_Skp2.
DR Pfam; PF00646; F-box; 1.

```

```

DR SMART; SMO0579; FBD; 1.
DR SMART; SMO0256; FBO; 1.
DR PROSITE; PS0181; FBO; 1.
DR LYASE.
SQ SEQUENCE 476 AA; 51900 MW; 43FF7DC3DD1305C CRC64;

Query Match      2.1%; Score 7; DB 10; Length 476;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 KTDKASE 40
Db 355 KTDKASE 361
|||||

RESULT 166
Q7VQ17 PRELIMINARY; PRT; 480 AA.
AC Q7VQ17
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE UNP-N-acetylmuramate:alanine ligase (EC 6.3.2.8).
GN MURC OR BFL143.
OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enderobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22784745; PubMed=12886019;
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis
of reduced genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393 (2003).
DR EMBL; BX248584; CAD83664.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 480 AA; 54194 MW; 2C8C6455543D369F CRC64;

Query Match      2.1%; Score 7; DB 16; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 MNLLRDK 265
Db 452 MNLLRDK 458
|||||

RESULT 167
Q9H677 PRELIMINARY; PRT; 485 AA.
AC Q9H677
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22531.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026184; BAB15387.1; -.
KW Hypothetical protein.
SQ SEQUENCE 485 AA; 54051 MW; 4474E549ACD560C3 CRC64;

Query Match      2.1%; Score 7; DB 4; Length 485;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 LLGELIL 237
Db 161 LLGELIL 167
|||||

RESULT 168
Q924K5 PRELIMINARY; PRT; 489 AA.
AC Q924K5
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Myocilin.
GN MYOC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Ahmed F., Tomarev S.I.;
RT "Rat Myoc/tigr gene: structure and changes in mRNA level in the retina
with induced ganglion cell damage";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289235; AAK83081.1; -.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF02191; OLF; 1.
DR SMART; SMO0284; OLF; 1.
SQ SEQUENCE 489 AA; 55094 MW; 68BD5BB9126E9006 CRC64;

Query Match      2.1%; Score 7; DB 11; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266
Db 130 NLLRDKS 136
|||||

RESULT 169
Q7Z6Q9 PRELIMINARY; PRT; 490 AA.
AC Q7Z6Q9
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DJ454G6.1 (Myocilin, trabecular meshwork inducible glucocorticoid
response (TIGR)).
GN MYOC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pavlitt R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98750; CAD92590.1; -.
SQ SEQUENCE 490 AA; 55336 MW; 70BCB267AB52428E CRC64;

Query Match      2.1%; Score 7; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266
Db 130 NLLRDKS 136
|||||

```

```

RESULT 170
ID Q866N2 PRELIMINARY; PRT; 490 AA.
AC Q866N2;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Myocilin.
GN MYOC.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Shepard A.R., Jacobson N., Sui H., Steely H.T., Lotery A.J.,
RA Stone E.M., Clark A.F.;
RT "Rabbit myocilin cloning: Implications for human myocilin
RT glycosylation and signal peptide usage.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY191317; AAC38666.1; -.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF02191; OLF; 1.
DR SMART; SM00284; OLF; 1.
SQ SEQUENCE 490 AA; 54895 MW; F975AC169461FD49 CRC64;

Query Match 2.1%; Score 7; DB 6; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266
Db 130 NLLRDKS 136
|||||

RESULT 171
ID Q924K4 PRELIMINARY; PRT; 490 AA.
AC Q924K4;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Myocilin.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ahmed F., Tomarev S.I.;
RT "Rat Myoc/Tigr gene: structure and changes in mRNA level in the retina
RT with induced ganglion cell damage.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289236; AAK83082.1; -.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF02191; OLF; 1.
DR SMART; SM00284; OLF; 1.
SQ SEQUENCE 490 AA; 55314 MW; 2B5C5475BC2A0425 CRC64;

Query Match 2.1%; Score 7; DB 11; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266
Db 130 NLLRDKS 136
|||||

RESULT 172
ID Q863A3 PRELIMINARY; PRT; 491 AA.
AC Q863A3;
DT 01-JUN-2003 (TRENBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Myocilin.
GN MYOC.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Fingert J.H., Clark A.F., Craig J.E., Alward W.L., Snibson G.R.,
RA Fingert J.H., Clark A.F., Craig J.E., Alward W.L., Snibson G.R.,
RA McLaughlin M., Tuttle L., Mackey D.A., Sheffield V.C., Stone E.M.;
RA McLaughlin M., Tuttle L., Mackey D.A., Sheffield V.C., Stone E.M.;
RT "Evaluation of the myocilin (MYOC) glaucoma gene in monkey and human
RT steroid-induced ocular hypertension.";
RL Invest. Ophthalmol. Vis. Sci. 42:145-152(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Fingert J.H., Clark A.F., Craig J.E., Alward W.L., Snibson G.R.,
RA McLaughlin M., Tuttle L., Mackey D.A., Sheffield V.C., Stone E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY190130; AA040254.1; -.
DR EMBL; AY190130; AA040254.1; JOINED.
DR EMBL; AY190129; AA040254.1; JOINED.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF02191; OLF; 1.
DR SMART; SM00284; OLF; 1.
SQ SEQUENCE 491 AA; 55513 MW; 3D18DE0BA4D6ABD CRC64;

Query Match 2.1%; Score 7; DB 6; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266
Db 131 NLLRDKS 137
|||||

RESULT 173
ID Q7SXV4 PRELIMINARY; PRT; 492 AA.
AC Q7SXV4;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=Body;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;

```

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]

RC SEQUENCE FROM N.A.  
 RP STRAIN=AB; TISSUE=Body;  
 RA Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC055232; AA05232.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 492 AA; 56252 MW; 4A2AD535B329AE78 CRC64;

Query Match 2.1%; Score 7; DB 13; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 SLOWKE 51  
 Db 282 SLOWAKE 288

## RESULT 174

Q85SN3 PRELIMINARY; PRT; 496 AA.

ID Q85SN3  
 AC Q85SN3  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative protein phosphatase 2C (Hypothetical protein).  
 GN O1003C07.6.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]

## SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;  
 RA McCombie W.R.; Spiegel L.; de la Bastide M.; Preston R.; Ferraro K.,  
 RA Kuit K.; Nascimben L.; Zlatavarn T.; Balija V.; Bell M.; Baker J.,  
 RA Santos L.; Miller B.; Katzenberger F.; Muller S.; King L.; Yang C.,  
 RA Dike S.; O'Shaughnessy A.; Palmer L.; Dedhia N.;  
 RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
 O1003C07, from chromosome 10, complete sequence."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]

## SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;  
 RA The Rice Chromosome 10 Sequencing Consortium;  
 RT "In-depth view of structure, activity, and evolution of rice  
 RT chromosome 10."  
 RL Science 300:1566-1569 (2003).  
 RN [3]

## SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;  
 RA Buell C.R.; Wing R.A.; McCombie W.R.; Messing J.; Yuan Q.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC113335; AA008826.1; -.  
 DR EMBL; AE017084; AAP53374.1; -.  
 DR Gramine; Q85SN3; -.

DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR InterPro; IPR001932; PP2C-like.  
 DR Pfam; PF00481; PP2C; 1.

DR SMART; SM00332; PP2C; 1.

DR SMART; SM00331; PP2C; 1.

KW Hypothetical protein.

SQ SEQUENCE 496 AA; 52120 MW; 133CB5C5CBA88753 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 496;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 VSKSLQA 48

Db 430 VSKSLQA 436

## RESULT 175

Q842J1 PRELIMINARY; PRT; 499 AA.

ID Q842J1  
 AC Q842J1  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Nucleoside triphosphate protein.  
 GN NPT.

OS Caedibacter caryophilus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Holosporaceae; Caedibacter.

OX NCBI\_TaxID=28906;

RN [1]

RP SEQUENCE FROM N.A.

RA Linka N.; Hurka H.; Lang F.; Burger G.; Winkler H.H.; Urbany C.,

RA Neuhaus E.;

RT "Phylogenetic analysis of non-mitochondrial nucleotide transport  
 RT proteins in intracellular bacteria and eukaryotes, plant plastids, and  
 RT identification of unique biochemical properties of the ATP/ADP-  
 RT transport protein from *Holospira obtusa* and *Caedibacter caryophila*  
 RT allowing an efficient energy parasitism."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ441310; CAD29686.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0005471; F:ATP/ADP antiporter activity; IEA.

DR InterPro; IPR004667; ADP\_ATP\_car.

DR Pfam; PF03219; TLC; 1.

DR TIGRFAMs; TIGR00769; AAA; 1.

SQ SEQUENCE 499 AA; 56443 MW; 49923A88A0F445EE CRC64;

Query Match 2.1%; Score 7; DB 2; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 GLLVTLI 83

Db 283 GLLVTLI 289

## RESULT 176

Q95VY0 PRELIMINARY; PRT; 500 AA.

ID Q95VY0  
 AC Q95VY0  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Skinny hedgehog (SD13634p).  
 GN RASP OR SIT OR CG11495.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA Chamoun Z.; Mann R.K.; Nellen D.; Bellotto M.; Beachy P.A.; Basler K.;

RT "Drosophila melanogaster skinny hedgehog (ski).";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Stapleton M.; Brokstein P.; Hong L.; Adbayani A.; Carlson J.,

RA Champe M.; Chavez C.; Dorsett V.; Dresnek D.; Farfan D.; Frise E.,

RA George R.; Gonzalez M.; Guarin H.; Kronmiller B.; Li P.; Liao G.,

RA Miranda A.; Mungall C.J.; Nunoo J.; Pacleb J.; Paragas V.; Park S.,

RA Patel S.; Phouanavong S.; Wan K.; Yu C.; Lewis S.E.; Rubin G.M.,

RA Celniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF398410; AAK97480.1; --  
DR EMBL; AY119202; AAM51062.1; --  
DR FLYBase; FBgn0024194; rasp.  
DR GO; GO:0007225; P:patched receptor ligand processing; IMP.  
DR InterPro; IPR004299; MBOAT\_fam.  
DR Pfam; PF03062; MBOAT; 1.  
SQ SEQUENCE 500 AA; 58148 MW; 5193D7823CB216B8 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 500;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LVTLLIA 84  
Db 124 LVTLLIA 130

RESULT 177

Q9VZU2 Q9VZU2 PRELIMINARY; PRT; 500 AA.  
AC Q9VZU2  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CG11495 protein (SIGHTLESS).  
GN RASP OR CG11495.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely.  
RX MEDLINE=2019606; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2140048; PubMed=11509241;  
RA Lee J.D., Treisman J.E.;  
RT "Sightless has homology to transmembrane acyltransferases and is  
RT required to generate active Hedgehog protein.";  
RL Curr. Biol. 11:1147-1152 (2001).  
DR EMBL; AE003477; AAF47725.1; --  
DR EMBL; AF393157; AAK73748.1; --  
DR FLYBase; FBgn0024194; rasp.  
DR GO; GO:0007225; P:patched receptor ligand processing; IMP.  
DR InterPro; IPR004299; MBOAT\_fam.  
DR Pfam; PF03062; MBOAT; 1.  
SQ SEQUENCE 500 AA; 58105 MW; 4498DC306976F2A2 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 500;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LVTLLIA 84  
Db 124 LVTLLIA 130

RESULT 178

Q9XWG6 Q9XWG6 PRELIMINARY; PRT; 502 AA.  
AC Q9XWG6; Q9U244;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-DRC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Y55D9A.2a protein.  
GN Y55D9A.2 OR Y55D9A.2A.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wallis J.M.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018 (1998).  
DR EMBL; AL032643; CAA21702.2; --  
DR FIR; T27190; T27190.  
DR WormPep; Y55D9A.2a; CE28712.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR000467; G\_patch.  
DR Pfam; PF01585; G\_patch; 1.  
DR SMART; SM00443; G\_patch; 1.  
DR PROSITE; PS50174; G\_PATCH; 1.  
SQ SEQUENCE 502 AA; 56846 MW; ACBDF9F6782F59D2 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 NEKEPPT 63  
Db 477 NEKEPPT 483

RESULT 179

Q95Q15 Q95Q15 PRELIMINARY; PRT; 511 AA.  
AC Q95Q15



DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Y55D9A.2b protein.  
 GN Y55D9A.2 OR Y55D9A.2B.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wallis J.M.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99069613; PubMed=9851916;  
 none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; AL032649; CAA21703.2; -;  
 DR PIR; T27191; T27191.  
 DR WormPep; Y55D9A.2b; CE28713.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR000467; G:patch.  
 DR Pfam; PF01585; G:patch; 1.  
 DR SMART; SM00443; G:patch; 1.  
 DR PROSITE; PS01174; G:PATCH; 1.  
 SQ SEQUENCE 511 AA; 57873 MW; 587CB67B8CBCE4F5 CRC64;  
  
 Query Match 2.1%; Score 7; DB 5; Length 511;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 57 NEKEPPT 63  
 Db 477 NEKEPPT 483  
  
 RESULT 180  
 Q9FZL2  
 ID Q9FZL2 PRELIMINARY; PRT; 520 AA.  
 AC Q9FZL2  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE F17L21.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,  
 RA Shinn P., Altafi H., Bei B., Chin C., Chio J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu J., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F17L21 from chromosome  
 RT I";  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chio J., Choi E., Lam B.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu K., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC004557; AAF99750.1; -;  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR008531; DUF813.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF05668; DUF813; 1.  
 DR PROSITE; PS00077; HEAT\_REPEAT; 1.  
 SQ SEQUENCE 520 AA; 56056 MW; 33F977D7766EE9CE CRC64;  
  
 Query Match 2.1%; Score 7; DB 10; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 301 LIEFLSS 307  
 Db 240 LIEFLSS 246  
  
 RESULT 181  
 Q9KND4  
 ID Q9KND4 PRELIMINARY; PRT; 521 AA.  
 AC Q9KND4  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Methyl-accepting chemotaxis protein.  
 GN VCA0031.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Seilers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae";  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004347; AAF95945.1; -;  
 DR PIR; B82508; B82508.  
 DR TIGR; VCA0031; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0006935; F:chemotaxis; IEA.  
 DR GO; GO:0007165; F:signal transduction; IEA.  
 DR InterPro; IPR004089; Chmtaxis\_transd.  
 DR InterPro; IPR003660; HAMP.  
 DR InterPro; IPR004090; Me\_chemotaxis.  
 DR Pfam; PF00672; HAMP; 1.

```

DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDCUR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS0111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 1.
DR Complete proteome.
SQ SEQUENCE 521 AA; 57258 MW; F86E09D49F2DA855 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LVLTLIA 84
Db 177 LVLTLIA 183

RESULT 182
Q8N6T0 PRELIMINARY; PRT; 522 AA.
AC Q8N6T0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to hypothetical protein FLJ22531.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028240; AH28240.1; -.
KW Hypothetical protein.
SQ SEQUENCE 522 AA; 58061 MW; C7A9100E2E10E9E1 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELIL 237
Db 161 LLGELIL 167

RESULT 183
Q8XE85 PRELIMINARY; PRT; 523 AA.
AC Q8XE85;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein z0521.
GN Z0521 OR ECS0472.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / BDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Fostel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533 (2001).

[2]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005221; AAG54768.1; -.
DR EMBL; AP002551; BAB33895.1; -.
DR PIR; D85538; D85538.
DR PIR; H90687; H90687.
DR InterPro; IPR001646; S-peptide repeat.
DR Pfam; PF00805; Pentapeptide_2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 523 AA; 60952 MW; 11E5D0F0FB4DD86 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 PLAKIIL 160
Db 356 PLAKIIL 362

RESULT 184
Q8LQ57 PRELIMINARY; PRT; 530 AA.
AC Q8LQ57;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P0702H08.17 protein.
GN P0702H08.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone: P0702H08."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003350; BAB91860.1; -.
DR Gramene; Q8LQ57; -.
SQ SEQUENCE 530 AA; 59292 MW; F8C866F1331B8A07 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQEL 73
Db 514 AQLAQEL 520

RESULT 185
Q9LG16 PRELIMINARY; PRT; 534 AA.
AC Q9LG16;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P0009G03.20 protein (P0030H07.3 protein).
GN P0009G03.20 OR P0030H07.3.
OS Oryza sativa (Rice).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P003G03";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P003H07";  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF002522; BAB03620.1; -;  
 DR EMBL; AF003045; BAB44041.1; -;  
 DR Gramene; O9LGI6; -;  
 DR InterPro; IPR002885; PPR.  
 DR InterPro; IPR008941; TPR-like.  
 DR Pfam; PF01535; PPR; 10.  
 DR TIGFAMS; TIGR00756; PPR; 10.  
 SQ SEQUENCE 534 AA; 61225 MW; 96CD7BC918119DD9 CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 534;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 EIVKILK 22  
 DB 85 EIVKILK 91  
 RESULT 186  
 Q9DW53 PRELIMINARY; PRT; 549 AA.  
 AC Q9DW53;  
 DT 01-MAR-2001 (TREMELrel. 16, Created)  
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
 DE Pr142.  
 GN R142.  
 OS Rat cytomegalovirus (strain Maastricht).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Muromegalovirus.  
 OX NCBI\_TaxID=79700;  
 RN [1]  
 RP SEQUENCE OF 1-201 FROM N.A.  
 RC STRAIN=Maastricht;  
 RX MEDLINE=20091365; PubMed=10623772;  
 RA Beisser P.S., Kloover J.S., Gauls G.E., Blok M.J., Bruggeman C.A.,  
 Vink C.;  
 RT "The 144 major histocompatibility complex class I-like gene of rat  
 cytomegalovirus is dispensable for both acute and long-term infection  
 in the immunocompromised host";  
 RL J. Virol. 74:1045-1050(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maastricht;  
 RX MEDLINE=20473137; PubMed=11018281;  
 RA Gruitjuijssen Y.K., Beuken E., Bruggeman C.A., Vink C.;  
 RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a  
 spliced transcript";

RL Virus Res. 69:119-130(2000).  
 DR EMBL; AF232689; AAF99237.1; -;  
 DR InterPro; IPR003360; US22.  
 DR Pfam; PF02393; US22; 1.  
 SQ SEQUENCE 549 AA; 61955 MW; D9B3680F32436D87 CRC64;  
 Query Match 2.1%; Score 7; DB 12; Length 549;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 52 ILCGTNE 58  
 DB 323 ILCGTNE 329  
 RESULT 187  
 O00781 PRELIMINARY; PRT; 551 AA.  
 AC O00781;  
 DT 01-JUL-1997 (TREMELrel. 04, Created)  
 DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Leishmania major.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LV-561;  
 RA Ortiz G., Segovia M.;  
 RT "Subtelomeric ORF";  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z93324; CAB07534.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 551 AA; 60874 MW; BC99686CD23A2F68 CRC64;  
 Query Match 2.1%; Score 7; DB 5; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 80 VTLIADL 86  
 DB 367 VTLIADL 373  
 RESULT 188  
 O23691 PRELIMINARY; PRT; 554 AA.  
 AC O23691;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN T19D16.24.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
 Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Wu D.,  
 Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RT "Arabidopsis thaliana chromosome 1 BAC T19D16 genomic sequence";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U95973; AAB65493.1; -;  
 DR PIR; F86244; F86244.  
 KW Hypothetical protein.  
 SQ SEQUENCE 554 AA; 60902 MW; 8188B41F9AE27BBA CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 554;

```
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AILEKQD 32
Db 161 AILEKQD 167

RESULT 189
Q20274 PRELIMINARY; PRT; 558 AA.
AC Q20274;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F41E7.2 protein.
GN F41E7.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68106; CAA92125.2; -
DR EMBL; Z49153; CAA92125.2; JOINED.
DR EMBL; Z49153; CAD30426.1; -
DR EMBL; Z68106; CAD30426.1; JOINED.
DR PIR; T22075; T22075.
DR WormPep; F41E7.2; CE30532.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015299; F:solute:hydrogen antiporter activity; IEA.
DR GO; GO:0006885; P:regulation of pH; IEA.
DR InterPro; IPR006153; Na_H porter.
DR Pfam; PF00999; Na_H_Exchange; 1.
SQ SEQUENCE 558 AA; 61977 MW; CCE9DD6E134CBB6E CRC64;

Query Match 2.1%; Score 7; DB 5; Length 558;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 PIVEILL 295
Db 457 PIVEILL 463

RESULT 190
Q8BK99 PRELIMINARY; PRT; 567 AA.
AC Q8BK99;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dynamin-like 120 kDa protein.
GN OPAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Eye;
MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK053861; BAC35561.1; -
DR MGD; MGI:1921393; Opal.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR InterPro; IPR001401; DYNAMIN.
DR Pfam; PF00350; dynamin; 1.
DR PRINTS; PR00195; DYNAMIN.
DR SMART; SM00053; DYNC; 1.
SQ SEQUENCE 567 AA; 64377 MW; CA9D6BC0FC7AAD52 CRC64;

Query Match 2.1%; Score 7; DB 11; Length 567;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELLIL 237
Db 210 LLGELLIL 216

RESULT 191
Q949V9 PRELIMINARY; PRT; 575 AA.
AC Q949V9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative auxin-responsive protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Barth J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Xu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashiraki Y., Ishida J., Jones T., Kamlya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F11C10.6/At2g46370 (GI:4559380).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050861; AAK92798.1; -
DR InterPro; IPR004993; GH3.
DR Pfam; PF03321; GH3; 1.
SQ SEQUENCE 575 AA; 64301 MW; A46BB58B2358D0B9 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 575;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKQ 298
Db 34 EILLKQ 40

RESULT 192
Q9SKE2 PRELIMINARY; PRT; 575 AA.
AC Q9SKE2; Q9LKI2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative auxin-responsive protein (FIN219).
GN AT2G46370 OR FIN219.
```



```
RT the African lion, Panthera leo. ";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385614; AAM46197.1; -.
DR GO; GO:0005739; C-mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . . ; IEA.
DR InterPro; IPR003916; NADH_oxrd5.
DR InterPro; IPR001750; Oxidored q1.
DR InterPro; IPR001516; Oxidored q1_N.
DR Pfam; PF00361; oxidored_q1_1_N.
DR Pfam; PF00682; oxidored_q1_N; 1.
DR PRINTS; PRO1434; NADHHDGNASES.
DR KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
SQ SEQUENCE 606 AA; 68282 MW; 1E23B01289E8E04B CRC64;

Query Match 2.1%; Score 7; DB 8; Length 606;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
DB 409 LLVTLIA 415

RESULT 196
Q973CO PRELIMINARY; PRT; 606 AA.
AC Q973CO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein ST0973.
GN ST0973.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kwarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo H. Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000984; BAB65993.1; -.
DR GO; GO:0004339; P:glucan 1,4-alpha-glucosidase activity; IEA.
DR GO; GO:0005976; P:polysaccharide metabolism; IEA.
DR InterPro; IPR00165; Glyco_hydro_15.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF00723; Glyco_hydro_15; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 606 AA; 70032 MW; F98C6DFB1446BD92 CRC64;

Query Match 2.1%; Score 7; DB 17; Length 606;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
DB 558 SLKLGE 564

RESULT 197
Q9DDA3 PRELIMINARY; PRT; 611 AA.
ID Q9DDA3
AC Q9DDA3;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RalB-binding protein (Fragment).
GN RLBP1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC Moreau J., Lebreton S.;
RT "Small G protein XRa1b cross-talks with Ras/Raf/MAPK pathway to
RT control the morphogenetic movement of gastrulation.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304845; CAC19674.1; -.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho GAP.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SMO0324; RhoGAP; 1.
DR PROSITE; PS50238; RHO GAP; 1.
DR NON TER 1
FT SEQUENCE 611 AA; 71592 MW; 2AE1A5D64F74B5C1 CRC64;

Query Match 2.1%; Score 7; DB 13; Length 611;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGT 111
DB 464 LRRQIGT 470

RESULT 198
Q9CS11 PRELIMINARY; PRT; 614 AA.
AC Q9CS11;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F6N7.4/AT5G52560.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 1; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RC Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RC Carninci P., Chen H., Cheuk R., Hayaahizaki Y., Ishida J., Jones T.,
RC Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RC Miranda M., Narusaka M., Nguyen M., Palm C.-J., Sakurai T., Satou M.,
RC Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RC Theologis A.;
RT "Full length cDNA of gene F6N7.4/AT5G52560 (GI:8953711).";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RC Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RC Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RC Chen H., Cheuk R., Hayaahizaki Y., Ishida J., Jones T., Kamiya A.,
RC Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RC Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RC Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RC Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F6N7.4/AT5G52560 (GI:8953711).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF360236; AAK25946.1; -.
DR EMBL; AY040035; AAK64093.1; -.

```

```

[1]
RP SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RP Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA "Arabidopsis cDNA clones.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057540; AAL09780.1; -
DR EMBL; AY140087; AAM98228.1; -
DR EMBL; BT008391; AAP37750.1; -
DR InterPro; IPR008938; ARM.
DR InterPro; IPR008531; DUF813.
DR InterPro; IPR000357; HEAT.
DR Pfam; PF05668; DUF813; 1
DR PROSITE; PS00077; HEAT_REPEAT; 1.
SQ SEQUENCE 625 AA; 67541 MW; 387B03D99CF9262B CRC64;

Query Match 2.1%; Score 7; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 LIEFLSS 307
Db 240 LIEFLSS 246
|||||
|||||

RESULT 201
Q98AT0 PRELIMINARY; PRT; 631 AA.
ID Q98AT0;
AC Q98AT0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter, ATP-binding protein, ExsA.
GN MLX5867.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyosaka C., Kohara M., Matsumoto M., Matsuoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

```



```

RT Mesorhizobium loti.
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003007; BAB52242.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP-binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 631 AA; 69489 MW; A7DAEBIP457A6985 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 631;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDAF 184
Db 133 DIASDAF 139

RESULT 202
ID Q81820 PRELIMINARY; PRT; 634 AA.
AC Q81820;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Heat shock protein 90 (Fragment).
GN HSP-90.
OS Bodo saliens.
OC Eukaryota; Euglenozoa; Kinetoplastida; Bodonidae; Bodo.
OX NCBI_TaxID=127146;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 50358;
RX MEDLINE=22337190; PubMed=12446799;
RA Simpson A.G.B., Lukes J., Roger A.J.;
RT "The Evolutionary History of Kinetoplastids and Their Kinetoplasts.";
RL Mol. Biol. Evol. 19:2071-2083(2002).
DR EMBL; AY122627; AAM93749.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003754; F:chaperone activity; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR001404; Hsp90.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00183; HSP90; 1.
DR PRINTS; PR00775; HEATSHOCK90.
DR SMART; SM00387; HATPase_c; 1.
FT NON_TER 1
FT NON_TER 634
SQ SEQUENCE 634 AA; 72719 MW; 02DF994CA541EAD6 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 634;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EKQDKKT 35
Db 218 EKQDKKT 224

RESULT 203
Q8KJ66 PRELIMINARY; PRT; 643 AA.
ID Q8KJ66

```

```

AC Q8KJ66;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable SACCHARIDE EXPORTING ABC transporter protein, ATP-binding and
DE permease domains.
GN EXSA.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R7A;
RX MEDLINE=21999272; PubMed=12003951;
RA Sullivan J.T., Trzebiatowski J.R., Cruickshank R.W., Gouzy J.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rossbach U.,
RA Stuart G.S., Weaver J.E., Webby R.J., de Bruijn F.J., Ronson C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A.";
RL J. Bacteriol. 184:3086-3095(2002).
DR EMBL; AL672114; CAD31371.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding.
SQ SEQUENCE 643 AA; 70759 MW; DA6253137BF605B5 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 643;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDAF 184
Db 145 DIASDAF 151

RESULT 204
Q7V9L8 PRELIMINARY; PRT; 647 AA.
ID Q7V9L8;
AC Q7V9L8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metallo-beta-lactamase superfamily hydrolase.
GN PRO1812.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / CCMP 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486;
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Ozta S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxyphototrophic genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).

```

DR EMBL; AE017166; AAQ00856.1; --  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 647 AA; 71209 MW; A46P67B9CE9FBAC CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 647;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 326 IKQIRD 332  
 Db 292 IKQIRD 298  
 RESULT 205  
 Q9LFX1 PRELIMINARY; PRT; 649 AA.  
 AC Q9LFX1  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE T7N9.27.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,  
 RA Shinn P., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC T7N9 from chromosome  
 RT 1.";  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC000348; AAF79870.1; --  
 DR InterPro; IPR008938; ARM  
 DR InterPro; IPR008531; DUF813.

DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF05668; DUF813; 1.  
 DR PROSITE; PS0077; HEAT\_REPEAT; 1.  
 SQ SEQUENCE 649 AA; 70051 MW; 40A16PF2F185CCCB CRC64;  
 Query Match 2.1%; Score 7; DB 10; Length 649;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 301 LIEFLSS 307  
 Db 240 LIEFLSS 246  
 RESULT 206  
 Q8KFR3 PRELIMINARY; PRT; 653 AA.  
 ID Q8KFR3  
 AC Q8KFR3  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ABC transporter, ATP-binding protein.  
 OS Chlorobium tepidum.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium.  
 OX NCBI\_TaxID=1097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=TLS / ATCC 49652 / DSM 12025;  
 RX MEDLINE=22103685; PubMed=12093901;  
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,  
 RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,  
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,  
 RA Holt I., Unayam L.A., Mason T., Brenner M., Shear T.P., Parksey D.,  
 RA Niemman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,  
 RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,  
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;  
 RT "The complete genome sequence of Chlorobium tepidum TLS, a  
 RT photosynthetic, anaerobic, green-sulfur bacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).  
 DR EMBL; AB012804; AAM71505.1; --  
 DR TIGR; CT0259; --  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0000166; P:nucleotide binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR001140; ABC\_TM transporter.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR Pfam; PF00664; ABC\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 653 AA; 73148 MW; 4A6DDC186EAF97EE CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 653;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 301 LIEFLSS 307  
 Db 273 LIEFLSS 279  
 RESULT 207  
 Q7V430 PRELIMINARY; PRT; 656 AA.  
 ID Q7V430  
 AC Q7V430

DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ABC transporter, multidrug efflux type.  
 GN FMT2143.  
 OS Prochlorococcus marinus (strain MIT 9313).  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=74547;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22825698; PubMed=12917642;  
 RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,  
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
 RA Webb E.A., Zinser E.R., Chisholm S.W.;  
 RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
 RT niche differentiation."  
 RL Nature 424:1042-1047(2003).  
 DR EMBL: BX572101; CAE22317.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 656 AA; 71066 MW; 25650D83488FFDA3 CRC64;  
 Query Match 2.1%; Score 7; DB 16; Length 656;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 105 LRRQIGT 111  
 Db 494 LRRQIGT 500  
 RESULT 208  
 Q80V96  
 ID Q80V96 PRELIMINARY; PRT; 658 AA.  
 AC Q80V96;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DE Hypothetical protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC049983; AA049983.1; -.  
 KW Hypothetical protein.  
 FT NON-ITER 1  
 SQ SEQUENCE 658 AA; 72605 MW; 7A3CECD9287071B0 CRC64;  
 Query Match 2.1%; Score 7; DB 11; Length 658;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 40 EEVSKSL 46  
 Db 596 EEVSKSL 602  
 RESULT 209  
 Q9VDX2  
 ID Q9VDX2 PRELIMINARY; PRT; 662 AA.  
 AC Q9VDX2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE CG5316-PB.  
 GN CG5316.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=107311132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA April J.F., Agayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foele C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Jbegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Paclet J., Faragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome."

```
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003726; AAF55666.2; -.
DR FlyBase; FBgn0038704; CG5316.
DR InterPro; IPR001310; HIT.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01230; HIT; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 2.
SQ SEQUENCE 662 AA; 76477 MW; D4DC720076FFB331 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 662;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 ISKPENL 255
Db 11 ISKPENL 17

RESULT 210
ID Q8MSG8 PRELIMINARY; PRT; 663 AA.
AC Q8MSG8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GM01362p.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118832; AAM50692.1; -.
DR FlyBase; FBgn0038704; CG5316.
DR InterPro; IPR001310; HIT.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01230; HIT; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 2.
SQ SEQUENCE 663 AA; 76528 MW; 00EC66A3728C73BC CRC64;

Query Match 2.1%; Score 7; DB 5; Length 663;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 ISKPENL 255
Db 11 ISKPENL 17

RESULT 211
ID Q8CCN1 PRELIMINARY; PRT; 673 AA.
AC Q8CCN1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PAAD and NACHT containing protein.
GN NAPL10 OR 6430548I20RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK032446; BAC27872.1; -.
DR MGD; MGI:2444084; Napl10.
DR GO; GO:0005351; F.sugar porter activity; IEA.
DR GO; GO:0009403; P.phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR InterPro; IPR002114; HPR_Serp_S.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAEIN; 1.
DR PROSITE; PS50837; NACHT; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
SQ SEQUENCE 673 AA; 76367 MW; EE773C592BEC7054 CRC64;

Query Match 2.1%; Score 7; DB 11; Length 673;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 191 LTRHKVL 197
Db 389 LTRHKVL 395

RESULT 212
Q9ZUE0
ID Q9ZUE0 PRELIMINARY; PRT; 731 AA.
AC Q9ZUE0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F508.10 protein.
GN F508.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Vysotskaia V.S.; Schwartz J.R., Yu G., Toriumi M., Liu S., Lenz C.,
RA Li J., Kremenetskaia I., Luros J., Altafi H., Gonzalez A., Araujo R.,
RA Buehler B., Conn L., Conway A.B., Dunn P., Hansen N., Huizar L.,
RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
RA Ecker J.R., Federpiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F508 sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Theologis A.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC005990; AAC98010.1; -.
DR PIR; B86369; B86369.
DR GO; GO:0005524; F.ATP binding; IEA.
```

DR GO: 0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO: 0016740; F:transferase activity; IEA.  
 DR GO: 0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR PROSITE: PS0107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 731 AA; 77639 MW; 45D93AD5C450001B CRC64;

Query Match 2.1%; Score 7; DB 10; Length 731;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 KVLVADF 201  
 |||||  
 Db 519 KVLVADF 525

RESULT 213  
 Q876G2 PRELIMINARY; PRT; 732 AA.  
 AC Q876G2  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE PH087 (Fragment).  
 OS Saccharomyces bayanus (Yeast) (Saccharomycetes uvarum).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4931;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=623-6C;  
 RA Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY14849; AAC32413.1; -.  
 DR GO: 0016020; C:membrane; IEA.  
 DR GO: 0016462; F:pyrophosphatase activity; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0008152; P:metabolism; IEA.  
 DR GO: 0006814; P:sodium ion transport; IEA.  
 DR InterPro: IPR000585; Hemopexin.  
 DR InterPro: IPR001899; Na/sul\_symport.  
 DR InterPro: IPR008162; Pyrophosphatase.  
 DR InterPro: IPR004331; SPX.  
 DR Pfam: PF00939; Na\_sulph\_symp; 1.  
 DR Pfam: PF03105; SPX; 1.  
 DR PROSITE: PS00024; HEMOPEXIN.  
 DR PROSITE: PS00387; PPASE; 1.  
 FT NON TER 1  
 SQ SEQUENCE 732 AA; 80605 MW; CB4DF258D567C314 CRC64;

Query Match 2.1%; Score 7; DB 3; Length 732;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SGLLVT 81  
 |||||  
 Db 593 SGLLVT 599

RESULT 214  
 Q8WZU0 PRELIMINARY; PRT; 734 AA.  
 AC Q8WZU0  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Related to SOK1 protein.  
 GN B8J24.140.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hohseisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL669990; CAD21116.1; -.  
 DR InterPro: IPR008862; Tcpl1.  
 DR Pfam: PF05794; Tcpl1; 1.  
 DR SQ SEQUENCE 734 AA; 81374 MW; 5B9F9A16C66359CF CRC64;

Query Match 2.1%; Score 7; DB 3; Length 734;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 AQELYS 76  
 |||||  
 Db 510 AQELYS 516

RESULT 215  
 Q9KQC3 PRELIMINARY; PRT; 758 AA.  
 AC Q9KQC3  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ferrous iron transport protein B.  
 GN VC2077.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae";  
 RL Nature 406:477-483 (2000).  
 DR EMBL; AE004281; AAF95223.1; -.  
 DR TIGR; B82122; B82122.  
 DR TIGR; VC2077; -.  
 DR GO: 0016020; C:membrane; IEA.  
 DR GO: 0015093; F:ferrous iron transporter activity; IEA.  
 DR GO: 0015684; F:ferrous iron transport; IEA.  
 DR InterPro: IPR003373; FeoB.  
 DR InterPro: IPR006073; GTP1\_OBG.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF02421; FeoB; 1.

```
DR PRINTS: PR00326; GTP10BG.
DR TIGRFRAMS; TIGR00437; feob; 1.
DR TIGRFRAMS; TIGR00231; small_GTP; 1.
KW Complete proteome.
SQ SEQUENCE 758 AA; 83114 MW; EC194EE2D28FEF28 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 758;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLLIAD 85
Db 332 LVTLLIAD 338
|||||

RESULT 216
Q9BSF5 PRELIMINARY; PRT; 777 AA.
AC Q9BSF5; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RW EMBL; BC05074; AA05074.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 777 AA; 86029 MW; 6BA69AB781A39956 CRC64;

Query Match 2.1%; Score 7; DB 4; Length 777;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EVSKSL 46
Db 715 EVSKSL 721
|||||

RESULT 217
Q8MXZ9 PRELIMINARY; PRT; 780 AA.
AC Q8MXZ9; 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Piwi-related protein.
GN TWI1.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RA Mochizuki K., Gorovskiy M.A.;
RT "Analysis of a piwi-related gene implicates small RNA in DNA rearrangement in Tetrahymena.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB084111; BAC02573.1; -.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
SQ SEQUENCE 780 AA; 89525 MW; 58D3FE686D6FE288 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 780;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
Db 97 SLKLGE 103
|||||

RESULT 218
Q8MQL1 PRELIMINARY; PRT; 780 AA.
AC Q8MQL1; 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CnJA protein.
GN CnJA.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84067828; PubMed=6646127;
RA Martindale D.W., Bruns P.J.;
RT "Cloning of abundant mRNA species present during conjugation of Tetrahymena thermophila: identification of mRNA species present exclusively during meiosis.";
RL Mol. Cell. Biol. 3:1857-1865(1983).
RN [2]
RP SEQUENCE FROM N.A.
RA Rosenauer A., Martindale D.W.;
RT "The characterization of cnJA, a Tetrahymena gene active only during meiosis.";
RL Thesis (1993), McGill University, Thesis Office.
RN [3]
RP SEQUENCE FROM N.A.
RA Martindale D.W., Rosenauer A.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV129082; AAM77972.2; -.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50821; PAZ; 1.
DR PROSITE; PS50822; PIWI; 1.
SQ SEQUENCE 780 AA; 89552 MW; 51DCDFB2E8AF1184 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 780;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
Db 97 SLKLGE 103
|||||

RESULT 219
Q8YNK6 PRELIMINARY; PRT; 786 AA.
AC Q8YNK6; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE WD-40 repeat-protein.
GN ALR4559.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shampo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003596; BAB76258.1; -.  
 DR FIC; AG2375; AG2375.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR SMART; SM00320; WD40; 7.  
 DR PROSITE; PS00678; WD REPEATS\_1; 1.  
 DR PROSITE; PS0082; WD REPEATS\_2; 5.  
 DR PROSITE; PS0294; WD REPEATS\_REGION; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 786 AA; 88363 MW; E054B1EF65628CBF CRC64;  
  
 Query Match 2.1%; Score 7; DB 16; Length 786;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 24 MIALEX 30  
 Db 282 MIALEX 288  
  
 RESULT 220  
 Q8TT76  
 ID Q8TT76 PRELIMINARY; PRT; 791 AA.  
 AC Q8TT76;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Hypothetical protein MA0561.  
 GN MA0561.  
 OS Methanosarcina acetivorans.  
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238;  
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
 RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
 RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,  
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,  
 RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of *Methanosarcina acetivorans* reveals extensive metabolic  
 RT and physiological diversity.";  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL; AE010717; AAM04005.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 791 AA; 90717 MW; 50BC61629A0CFE57 CRC64;  
  
 Query Match 2.1%; Score 7; DB 17; Length 791;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 36 DKASEEV 42  
 Db 222 DKASEEV 228

RESULT 221  
 Q8UCC3  
 ID Q8UCC3 PRELIMINARY; PRT; 818 AA.  
 AC Q8UCC3;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Mannosidase.  
 GN ATU2575 OR AGR\_C 4665.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OX NCBI\_TaxID=176239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21608550; PubMed=11743193;  
 RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L.,  
 RA Kura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 RT C58.";  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21608551; PubMed=11743194;  
 RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmieu K., Gordon J., Vaudin M., Tartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT *Agrobacterium tumefaciens* C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009203; AAL43556.1; -.  
 DR EMBL; AE008170; AAK88298.1; -.  
 DR FIC; A97668; A97668.  
 DR PIR; AF2892; AF2892.  
 DR InterPro; IPR008973; Gal\_bind\_like.  
 KW Complete proteome.  
 SQ SEQUENCE 818 AA; 90067 MW; 06A011D8EE9D234E CRC64;  
  
 Query Match 2.1%; Score 7; DB 16; Length 818;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 230 KILGELI 236  
 Db 239 KILGELI 245  
  
 RESULT 222  
 O66933  
 ID O66933 PRELIMINARY; PRT; 831 AA.  
 AC O66933;  
 DT 01-AUG-1998 (TRENBLrel. 07, Created)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Mannose-1-phosphate guanylttransferase.  
 GN MGS OR AQ 718.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
DR EMBL; AE000704; AAC06893.1; -.
DR PIR; A70363; A70363.
DR GO; GO:0016868; F:intramolecular transferase activity, phospho. . .; IEA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001451; Hexapep transf.
DR InterPro; IPR005835; NTP transferase.
DR InterPro; IPR005844; PG_PWW_ABAI.
DR Pfam; PF00132; hexapep_4.
DR Pfam; PF00483; NTP transferase; 1.
DR Pfam; PF02878; PGM_PWW_I; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 831 AA; 93607 MW; F39513DC53594EA6 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 831;
Best Local Similarity 100.0%; Pred. No. 5.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 219 SENYVTK 225
Db 533 SENYVTK 539
|||||

RESULT 223
Q8ZE91 PRELIMINARY; PRT; 846 AA.
ID Q8ZE91
AC Q8ZE91
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative virulence factor.
GN YPO2291.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;

[1]
SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaiha M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414151; CAC91096.1; -.
DR PIR; AD0279; AD0279.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001005; Myb_DNA_binding.
DR PROSITE; PS00334; MYB_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 846 AA; 95345 MW; DFB8859195F1C4C7 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 846;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILLK 296
Db 564 IVEILLK 570
|||||

RESULT 225
Q8XIW9 PRELIMINARY; PRT; 866 AA.
ID Q8XIW9
AC Q8XIW9
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DNA polymerase I.
GN POLA OR CPE1994.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;

[1]
SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003192; BAB81700.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008409; F:5'-3' exonuclease activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
```

```
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR008918; 5_3_exo_C.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo N I.
DR InterPro: IPR003153; HHH_1.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuc_N; 1.
DR Pfam: PF00476; DNA_pol_A; 1.
DR PRINTS; PR00868; DNAPOLI.
DR SMART; SM00475; 53EXOG; 1.
DR SMART; SM00278; Hhh1; 1.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00482; POLAC; 1.
DR TIGR; TIGR00593; Pola; 1.
DR TIGR; TIGR00593; Pola; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Complete proteome.
SQ SEQUENCE 866 AA; 99118 MW; E7F786F720146615 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 866;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 DEKNYLI 326
DB 344 DEKNYLI 350
|||||

RESULT 226
Q9H706 PRELIMINARY; PRT; 875 AA.
AC Q9H706
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ21610.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025263; BAB15094.1; -.
DR EMBL; AK025263; BAB15094.1; -.
KW Hypothetical protein.
SQ SEQUENCE 875 AA; 97099 MW; 10A0E9147A9BFCB1 CRC64;

Query Match 2.1%; Score 7; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EEVSKSL 46
DB 813 EEVSKSL 819
|||||

RESULT 227
Q8ND03 PRELIMINARY; PRT; 877 AA.
AC Q8ND03
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (fragment).
GN DKFZP434N0931.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Weil B., Wiemann S.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834491; CAD39149.1; -.
DR EMBL; AL834491; CAD39149.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 877 AA; 97314 MW; 95630B42D96FC29D CRC64;

Query Match 2.1%; Score 7; DB 4; Length 877;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EEVSKSL 46
DB 815 EEVSKSL 821
|||||

RESULT 228
Q883V3 PRELIMINARY; PRT; 887 AA.
AC Q883V3
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensor protein KdpD.
GN KDPD OR PSPT02245.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Daviden T.,
RA White O., Fraser C., Collmer A.;
RA "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS016963; AAO55761.1; -.
DR EMBL; AS016963; AAO55761.1; -.
DR TIGR; PSPT02245; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006950; P:response to stress; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR003852; KdpD.
DR InterPro; IPR006016; Usp_dom.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HSKA; 1.
DR Pfam; PF02702; KdpD; 1.
DR Pfam; PF00382; Usp; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PROSITE; PS0109; HIS_KIN; 1.
KW Complete proteome.
SQ SEQUENCE 887 AA; 96814 MW; B45A5C89EB8B6D69 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 887;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQL 73
DB 635 AQLAQL 641
|||||
```

```
RESULT 229
007686 PRELIMINARY; PRT; 902 AA.
ID O07686
AC O07686;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE ORF A protein.
OS Listeria seeligeri.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1640;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLCC3379;
RA Lampidis R., Krefit J.;
RT "plcA/PrfA operon of Listeria seeligeri.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X97014; CAA5738.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000437; Prok_lipoProt_S.
DR TIGRFAMs; TIGR01167; LPXFG anchor; 1.
DR PROSITE; PS00013; PROKAR_LIPOPEPTIDIN; 1.
SQ SEQUENCE 902 AA; 99415 MW; 8DADD11EF7C056D1 CRC64;

Query Match 2.1%; Score 7; DB 2; Length 902;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILKD 23
DB 684 IVKILKD 690

RESULT 230
Q875W4 PRELIMINARY; PRT; 916 AA.
ID Q875W4
AC Q875W4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PHO87.
OS Saccharomyces castellii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=27288;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS4309;
RX MEDLINE=22482865; PubMed=12594514;
RA Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;
RT "Yeast genome duplication was followed by asynchronous differentiation of duplicated genes.";
RL Nature 421:848-852(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS4309;
RA Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY144951; AAC032514.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_sympot.
DR InterPro; IPR004331; SPX.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR Pfam; PF03105; SPX; 1.
SQ SEQUENCE 916 AA; 101476 MW; B18AC6008B6F69A6 CRC64;

Query Match 2.1%; Score 7; DB 3; Length 916;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 231
Q9FSES PRELIMINARY; PRT; 922 AA.
ID Q9FSES
AC Q9FSES;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Lipoxigenase (EC 1.13.11.12).
GN RCI-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kusabue; TISSUE=Leaf;
RX MEDLINE=20453114; PubMed=10998053;
RA Schaffrath U., Zabbai F., Dudler R.;
RT "Characterization of RCI-1, a chloroplastic rice lipoxigenase whose synthesis is induced by chemical plant resistance activators.";
RL Eur. J. Biochem. 267:5935-5942(2000).
DR EMBL; AJ270938; CAC01439.1; -.
DR HSSP; P08170; 2SBL.
DR Gramene; Q9FSES; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016165; F:lipoxigenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000907; Lipoxigenase.
DR InterPro; IPR008976; PLAT LH2.
DR Pfam; PF0305; lipoxigenase; 1.
DR PRINTS; PR00087; LIPOXYGENASE.
DR PROSITE; PS00711; LIPOXYGENASE_1; 1.
DR PROSITE; PS00081; LIPOXYGENASE_2; 1.
KW Oxidoreductase.
SQ SEQUENCE 922 AA; 104687 MW; B0DD0C172A2DFAE0 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 922;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLITRHK 195
DB 64 DLITRHK 70

RESULT 232
Q88AL5 PRELIMINARY; PRT; 1053 AA.
ID Q88AL5
AC Q88AL5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Cation efflux family protein.
GN PSPT00375.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Felblyum T., Gwin M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay R., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
```

```
RA White O., Fraser C., Collmer A.;  
RT "Complete sequence of Pseudomonas syringae."  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A8016857; AAO53919.1; -  
DR TIGR; PSPT00375; -  
DR InterPro; IPR001036; Acrflvin_res.  
DR Pfam; PF00873; ACR_tran; 1.  
DR PRINTS; PR00702; ACRIFLAVINRP.  
KW Complete proteome  
SQ SEQUENCE 1053 AA; 113670 MW; EDB6A557CE363D71 CRC64;  
  
Query Match 2.1%; Score 7; DB 16; Length 1053;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 81 TLIAADIQ 87  
Db 647 TLIAADIQ 653  
|||||  
|  
  
RESULT 233  
Q9A9P9 PRELIMINARY; PRT; 1055 AA.  
AC Q9A9P9;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE OmpA-related protein.  
GN CC0925.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
OC Caulobacteraceae; Caulobacter.  
OX NCBI_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Ely B.,  
RA Ptočka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA DeBoy R.J., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus."  
RL PROC. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL; A8005769; AAK22909.1; -  
DR PIR; A87364; A87364.  
DR TIGR; CC0925; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:004872; F:receptor activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR008989; Carboxypep_reg.  
DR InterPro; IPR000531; TonB_boxC.  
DR Pfam; PF00593; TonB_dep_Reg; 1.  
KW Complete proteome.  
SQ SEQUENCE 1055 AA; 112904 MW; F31D4E15E0C8CD78 CRC64;  
  
Query Match 2.1%; Score 7; DB 16; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 262 LRDKSPN 268  
Db 285 LRDKSPN 291  
|||||  
|  
  
RESULT 234  
Q8TAB3 PRELIMINARY; PRT; 1094 AA.  
ID Q8TAB3  
AC Q8TAB3  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE BA99E24.1.1 (Protocadherin 19 (Hypothetical protein KIAA1313))  
DE (fragment).  
GN PCDH19.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chapman J.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL355593; CAD24084.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005194; F:cell adhesion molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.  
DR InterPro; IPR002126; Cadherin.  
DR Pfam; PF00028; cadherin; 6.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM0112; CA; 6.  
DR PROSITE; PS00232; CADHERIN_1; 5.  
DR PROSITE; PS00268; CADHERIN_2; 6.  
KW Hypothetical protein; Calcium; Calcium-binding; Cell adhesion;  
KW Glycoprotein.  
FT NON_TER 1  
SQ SEQUENCE 1094 AA; 120316 MW; EFD59CF96F2EFFE CRC64;  
  
Query Match 2.1%; Score 7; DB 4; Length 1094;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 75 SSGLLVT 81  
Db 70 SSGLLVT 76  
|||||  
|  
  
RESULT 235  
Q27764 PRELIMINARY; PRT; 1115 AA.  
ID Q27764  
AC Q27764;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ATPase.  
GN YEL6.  
OS Plasmodium yoelii.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI_TaxID=5861;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91161669; PubMed=2150071;  
RA Murakami K., Tanabe K., Takada S.;  
RT "Structure of a Plasmodium yoelii gene-encoded protein homologous to  
the Ca(2+)-ATPase of rabbit skeletal muscle sarcoplasmic reticulum."  
RL J. Cell Sci. 97:487-495(1990).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASE FAMILY (E1-E2  
ATPASES).  
DR EMBL; X55197; CAA38982.1; -  
DR PIR; A45761; A45761.  
DR HSP; P04191; LEUL.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0005388; F:calcium-transporting ATPase activity; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0006816; P:calcium ion transport; IEA.  
DR GO; GO:0006812; P:cation transport; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0015992; P:proton transport; IEA.  
DR InterPro; IPR001757; ATPase_E1-E2.
```

```

DR InterPro; IPR005782; Calcium ATPase.
DR InterPro; IPR006068; Cation ATPase C.
DR InterPro; IPR004014; Cation ATPase N.
DR InterPro; IPR008250; E1-E2 ATPase_reg.
DR InterPro; IPR005834; Hydrolyase.
DR InterPro; IPR000695; H ATPase.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGR; TIGR01116; ATPase-IIA1_Ca; 1.
DR TIGR; TIGR01494; ATPase_P-type; 8.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
KW ATP-binding; Hydrolyase; Phosphorylation; Transmembrane.
SQ SEQUENCE 1115 AA; 126718 MW; 689698773DBE49B6 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 1115;
Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILKD 23
Db 784 IVKILKD 790

RESULT 236
QX0R4 PRELIMINARY; PRT; 1170 AA.
AC Q9X0R4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chromosome segregation SMC protein, putative.
GS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Uitterback T.K., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AB001774; AAD36257.1; -.
DR PIR; A72287; A72287.
DR PDB; 1GXI; 09-MAY-02.
DR TIGR; TM1182; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0007059; P:chromosome segregation; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF02463; SMC_N; 1.
DR Pfam; PF02463; SMC_N; 1.
DR TIGR; TIGR00650; MG442; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
KW Complete proteome.
SQ SEQUENCE 1170 AA; 137587 MW; 81B203B54B336DE4 CRC64;

Query Match 2.1%; Score 7; DB 16; Length 1170;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 KYVELST 176
Db 310 KYVELST 316

RESULT 237
Q803P9 PRELIMINARY; PRT; 1189 AA.
AC Q803P9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to kinectin 1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044389; AAH44389.1; -.
SQ SEQUENCE 1189 AA; 134074 MW; 1D1838616CC21928 CRC64;

Query Match 2.1%; Score 7; DB 13; Length 1189;
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LRDKSPN 268
Db 359 LRDKSPN 365

RESULT 238
Q86J27 PRELIMINARY; PRT; 1243 AA.
AC Q86J27;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115608; AAO51792.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1243 AA; 147185 MW; 32638DF439FB585C CRC64;

Query Match 2.1%; Score 7; DB 5; Length 1243;
Best Local Similarity 100.0%; Pred. No. 7.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307

```

```

Db      501 LIEFLSS 507
|||||
RESULT 239
Q86925 PRELIMINARY; PRT; 1244 AA.
AC Q86925;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein 2.
OS Aura virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=44158;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266268; PubMed=7747434;
RA Rumenapf T.H.;
RT "Aura virus is a New World representative of Sindbis-like viruses.";
RL Virology 208:621-633(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Rumenapf T.H.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF126284; AAD13623.1; -.
DR HSSP; P03316; IKXF.
DR MEROPS; C09.001; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019028; C:viral envelope; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000930; Peptidase_S3.
DR Pfam; PF00944; Alpha_core; 1.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
SQ SEQUENCE 1244 AA; 137116 MW; F86B682234ED2F46 CRC64;

Query Match 2.1%; Score 7; DB 12; Length 1244;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 KYVELST 176
Db 912 KYVELST 918
|||||
RESULT 240
Q62644 PRELIMINARY; PRT; 1250 AA.
AC Q62644;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE N-methyl-D-aspartate receptor NMDAR2C subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 2.1%; Score 7; DB 10; Length 1280;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
Db 493 SLKLGE 499
|||||
STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Boulter J., Pecht G.;
RT "Nucleotide Sequence of rat NMDA receptor subunit gene NMDAR2C.";
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR ENBL; U08259; AAA17832.1; -.
DR PIR; B45219; B45219.
DR HSSP; P19491; IGR2.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005234; F:glutamate-gated ion channel activity; IEA.
DR GO; GO:0004970; F:ionotropic glutamate receptor activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR001320; Ion_glu_receptor.
DR InterPro; IPR001508; NMDA_receptor.
DR InterPro; IPR001311; SBP/glu_receptor.
DR Pfam; PF00060; lig_chan; 1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PEPF; 1.
KW Receptor.
SQ SEQUENCE 1250 AA; 136706 MW; D45A8BF27B2A79BE CRC64;

Query Match 2.1%; Score 7; DB 11; Length 1250;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEVAQL 69
Db 114 TEVAQL 120
|||||
RESULT 241
Q9FTA4 PRELIMINARY; PRT; 1280 AA.
ID Q9FTA4;
AC Q9FTA4;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
GN F8L15.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buysbaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielens J., Van Montagu M., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; ALJ92174; CAC08331.1; -.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF01535; PPR; 26.
DR TIGRPFAMs; TIGR00756; PPR; 23.
KW Hypothetical protein.
SQ SEQUENCE 1280 AA; 144481 MW; ED59CE0B9EA89E43 CRC64;

Query Match 2.1%; Score 7; DB 10; Length 1280;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234
Db 493 SLKLGE 499
|||||

```

```

KW Receptor.
SQ SEQUENCE 1303 AA; 153737 MW; 5B122456F4FBEP9D CRC64;

Query Match      2.1%; Score 7; DB 5; Length 1303;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307
DB 611 LIEFLSS 617

RESULT 244
Q86J26 PRELIMINARY; PRT; 1297 AA.
ID AC Q86J26;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115608; AA051793.1; -
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008271; Ser thr pkin AS.
DR PROSITE; PS00108; PROTEIN_KINASE_S7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1297 AA; 152631 MW; 47F8435D354A5297 CRC64;

Query Match      2.1%; Score 7; DB 5; Length 1297;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307
DB 532 LIEFLSS 538

RESULT 243
Q86HY0 PRELIMINARY; PRT; 1303 AA.
ID AC Q86HY0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to G-protein-coupled receptor at plasma membrane; interactions
DE in two-hybrid system with Gpa2p; Gprlp.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115681; AA052391.1; -
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR002110; ANK.

KW Receptor.
SQ SEQUENCE 1303 AA; 153737 MW; 5B122456F4FBEP9D CRC64;

Query Match      2.1%; Score 7; DB 5; Length 1303;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307
DB 611 LIEFLSS 617

RESULT 245
Q9SL02 PRELIMINARY; PRT; 1316 AA.
ID AC Q9SL02; Q9M6P9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative RAD50 DNA repair protein (DNA repair-recombination protein)
DE (At2g31970/F22D22.28).
GN AT2G31970 OR RAD50.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]

```





Q8JHV7 PRELIMINARY; PRT; 1785 AA.  
 ID Q8JHV7  
 AC Q8JHV7  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 22, Last annotation update)  
 DE Laminin beta 1.  
 GN LAMB1.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955.  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=12070089;  
 RX MEDLINE=22065263; PubMed=12070089;  
 RA Parsons M.J., Pollard S.M., Saude L., Feldman B., Coutinho P.,  
 RA Hirst E.M., Stemple D.L.;  
 RT "Zebrafish mutants identify an essential role for laminins in  
 RT notochord formation.";  
 RL Development 129:3137-3146(2002).  
 DR EMBL: AF468049; AM61767.1;  
 DR GO: GO:0005578; C:extracellular matrix; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR008211; Laminin\_EGF.  
 DR Pfam: PF00053; laminin\_EGF; 13.  
 DR Pfam: PF00055; laminin\_Nterm; 1.  
 DR PRINTS: PR00011; EGF\_LAMININ.  
 DR SMART: SM00180; EGF\_Lam; 13.  
 DR SMART: SM00136; LamNT; 1.  
 DR PROSITE: PS00022; EGF\_1; 10.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01246; LAMININ\_TYPE\_EGF; 12.  
 KW Laminin EGF-like domain.  
 SQ SEQUENCE 1785 AA; 197279 MW; 1BD338A0B6AB9F26 CRC64;

Query Match 2.1%; Score 7; DB 13; Length 1785;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 325 LIKQIRD 331  
 Db 1501 LIKQIRD 1507

RESULT 249  
 Q9VMJ7 PRELIMINARY; PRT; 1838 AA.  
 AC Q9VMJ7  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG9088 protein (ID40310p).  
 GN LID OR CG9088.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=10731132;  
 RX STRAIN=Berkley;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwack C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celinker S.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB003612; AAF52319.1; -.  
 DR EMBL: AY095051; AAM11379.1; -.  
 DR FlyBase: FBgn0031759; lld.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001606; ARID.  
 DR InterPro: IPR003347; TF\_JmJC.  
 DR InterPro: IPR003349; TF\_JmJN.  
 DR InterPro: IPR004198; Znf\_C5HC2.  
 DR InterPro: IPR001965; Znf\_PHD.  
 DR Pfam: PF01388; ARID; 1.  
 DR Pfam: PF02373; JmJC; 1.  
 DR Pfam: PF02375; JmJN; 1.  
 DR Pfam: PF00628; PHD; 3.  
 DR Pfam: PF02928; zf-C5HC2; 1.  
 DR SMART: SM00501; BRIGHT; 1.  
 DR SMART: SM00558; JmJC; 1.  
 DR SMART: SM00545; JmJN; 1.  
 DR SMART: SM00249; PHD; 3.  
 DR PROSITE: PS01359; ZF\_PHD 1; 2.  
 DR PROSITE: PS00016; ZF\_PHD 2; 3.  
 SQ SEQUENCE 1838 AA; 203591 MW; E01BDB89027F9F50 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 1838;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 KNPAEIV 18  
 Db 1258 KNPAEIV 1264

```

RESULT 250
Q86IV6
ID Q86IV6 PRELIMINARY; PRT: 1838 AA.
AC Q86IV6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OK NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116986; AA051900.1; -.
DR GO; GO:0004866; F:endorpeptidase.inhibitor activity; IEA.
DR InterPro; IPR002160; Kunitz_legume.
DR PROSITE; PS00283; SOYBEAN_KUNITZ; 1.
KW Hypothetical protein.
SQ SEQUENCE 1838 AA; 208417 MW; 8E5F1B2533273199 CRC64;

Query Match 2.1%; Score 7; DB 5; Length 1838;
Best Local Similarity 100.0%; Pred.No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 VTQIFNN 103
Db 611 VTQIFNN 617

```

Search completed: April 12, 2004, 10:36:50  
Job time : 60 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:28:43 ; Search time 59 Seconds

(without alignments)  
1613.873 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 337

Sequence: 1 MKKNPLPSKSHNPAEIVKI.....FADEKNVLIKQIRDKKTAP 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : A Geneseq 29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	337	100.0	337	3	AAY94247 Human cal
2	337	100.0	337	4	AAB82090 Human acu
3	337	100.0	337	4	AM39078 Human pol
4	227	67.4	237	4	AM40864 Human pol
5	188	55.8	289	4	AAB94139 Human pro
6	23	6.8	341	3	AAY94248 Human cal
7	23	6.8	341	4	AAB48970 Human ANI
8	23	6.8	350	4	AAB20387 Human acu
9	23	6.8	496	4	AAE10858 Galt-huma
10	23	6.8	552	4	AAE10859 Lexa-huma
11	17	5.0	354	4	ABG23844 Novel hum
12	15	4.5	339	3	AAY94249 Drosophil
13	15	4.5	339	4	ABE60392 Drosophil
14	14	4.2	377	3	AAY94250 C. eleg
15	12	3.6	383	4	ABG23843 Novel hum
16	12	3.6	639	4	ABG25372 Novel hum
17	8	2.4	139	3	AAQ41153 Zea mays
18	8	2.4	148	3	AAQ41152 Zea mays
19	8	2.4	154	3	AAQ41151 Zea mays
20	8	2.4	464	4	ABE59571 Drosophil
21	8	2.4	923	4	ABE58067 Drosophil
22	7	2.1	44	4	ABG03535 Novel hum
23	7	2.1	46	3	AAQ55362 Arabidops
24	7	2.1	51	4	ABE53042 Escherich
25	7	2.1	83	5	ABP05605 Human ORF

26	7	2.1	97	5	ABP08677 Human ORF
27	7	2.1	98	4	AM86795 Human imm
28	7	2.1	99	4	AAQ89310 Human sec
29	7	2.1	99	6	ABP76204 Human GEN
30	7	2.1	99	6	ABP76061 Human GEN
31	7	2.1	107	4	AAO09956 Human pol
32	7	2.1	113	4	AAQ74439 Human col
33	7	2.1	115	6	ABM71944 Staphyloc
34	7	2.1	139	4	ABE65844 Drosophil
35	7	2.1	141	6	ABU31422 Protein e
36	7	2.1	176	2	AAQ29337 Human sec
37	7	2.1	176	4	AAU39088 Human sec
38	7	2.1	176	5	ABE55797 Human pol
39	7	2.1	192	4	AAU31106 Novel hum
40	7	2.1	197	3	AAQ05091 Arabidops
41	7	2.1	199	7	ADC31499 Human nov
42	7	2.1	205	2	AAW11481 D. immiti
43	7	2.1	213	3	AAQ30715 Arabidops
44	7	2.1	213	3	AAQ23888 Arabidops
45	7	2.1	213	3	AAQ45275 Arabidops
46	7	2.1	215	4	AAU25691 G protein
47	7	2.1	217	6	ADA54954 Human pro
48	7	2.1	221	2	AAW11479 D. immiti
49	7	2.1	236	5	ABP02921 Human ORF
50	7	2.1	245	5	AAE24494 Human RAT
51	7	2.1	257	4	AAU63772 Propionib
52	7	2.1	257	6	ABM60291 Propionib
53	7	2.1	263	3	AAW41773 Human ORF
54	7	2.1	275	5	ABU05677 M. tuberc
55	7	2.1	282	6	ABU24179 Protein e
56	7	2.1	290	3	AAQ23887 Arabidops
57	7	2.1	291	2	AAQ48312 Human pro
58	7	2.1	300	3	AAQ45274 Arabidops
59	7	2.1	300	3	AAQ30714 Arabidops
60	7	2.1	302	6	ABM71583 Staphyloc
61	7	2.1	305	3	AAQ30713 Arabidops
62	7	2.1	308	3	AAQ73387 HTSM clon
63	7	2.1	320	3	AAQ05090 Arabidops
64	7	2.1	326	3	AAQ51053 Arabidops
65	7	2.1	332	2	AAQ17865 Sulfolobu
66	7	2.1	332	2	AAQ17866 Sulfolobu
67	7	2.1	337	4	ABG27865 Novel hum
68	7	2.1	340	5	ABE97360 Novel hum
69	7	2.1	343	3	AAQ45273 Arabidops
70	7	2.1	343	6	ADA32843 Acinetoba
71	7	2.1	344	6	ABU23650 Protein e
72	7	2.1	345	3	AAQ05089 Arabidops
73	7	2.1	368	5	ABE91447 Herbicida
74	7	2.1	374	4	AAQ73684 Human col
75	7	2.1	377	3	AAQ06308 Arabidops
76	7	2.1	377	5	ABE92224 Herbicida
77	7	2.1	384	4	AAE85255 Plant por
78	7	2.1	400	3	AAQ51052 Arabidops
79	7	2.1	422	3	AAQ39773 Arabidops
80	7	2.1	422	5	AAE24493 Human RAT
81	7	2.1	422	6	ABE83334 Human NOV
82	7	2.1	432	6	ABU31788 Protein e
83	7	2.1	439	6	ABE52729 Escherich
84	7	2.1	452	4	ABP73473 Candida a
85	7	2.1	458	5	AAQ32191 Arabidops
86	7	2.1	482	3	AAQ32191 Arabidops
87	7	2.1	490	3	AAQ49289 Mouse GLC
88	7	2.1	492	2	AAW37733 Cytochrom
89	7	2.1	492	2	AAW41159 Zebrafish
90	7	2.1	492	4	AAE85155 Zebrafish
91	7	2.1	492	5	ABP52146 Zebrafish
92	7	2.1	492	5	AAE15325 Zebrafish
93	7	2.1	497	2	AAE98225 Trabecula
94	7	2.1	497	2	AAW61390 Glucocort
95	7	2.1	497	3	AAQ39772 Arabidops
96	7	2.1	500	4	ABE65679 Drosophil
97	7	2.1	500	6	ABU07341 Human TIG
98	7	2.1	502	7	AAE54515 Rat Prote

99	7	2.1	504	2	AAW64669	AAW64669 Human TIG	172	6	1.8	15	2	AAW07418	AAW07418 HSV-1 TK
100	7	2.1	504	2	AAW60670	AAW60670 Human glia	173	6	1.8	15	2	AAW97143	AAW97143 HSV-1 TK
101	7	2.1	504	2	AAW70496	AAW70496 Trabecula	174	6	1.8	15	2	AAW97150	AAW97150 HSV-1 TK
102	7	2.1	504	2	AAW07393	AAW07393 Human TIG	175	6	1.8	15	6	ABG74902	ABG74902 Human TFR
103	7	2.1	504	2	AAW73500	AAW73500 Trabecula	176	6	1.8	15	6	ABG99202	ABG99202 Thymidine
104	7	2.1	504	2	AAW89391	AAW89391 Human tra	177	6	1.8	15	6	ABG99209	ABG99209 Thymidine
105	7	2.1	504	3	AAW49288	AAW49288 Human GLC	178	6	1.8	18	4	AAW22114	AAW22114 Peptide #
106	7	2.1	504	3	AAW93971	AAW93971 A trabecu	179	6	1.8	18	4	ABW44520	ABW44520 Peptide #
107	7	2.1	504	3	AAW51051	AAW51051 Arabidops	180	6	1.8	18	4	AAW38567	AAW38567 Peptide #
108	7	2.1	504	4	AAW03744	AAW03744 Human tra	181	6	1.8	18	4	AAW78321	AAW78321 Human bon
109	7	2.1	504	4	AAU09184	AAU09184 Human PRO	182	6	1.8	18	4	AAW65705	AAW65705 Human bra
110	7	2.1	504	4	AAW48845	AAW48845 Human TIG	183	6	1.8	18	4	ABG59929	ABG59929 Human liv
111	7	2.1	504	6	ABG75692	ABG75692 Human tra	184	6	1.8	18	5	ABG47346	ABG47346 Human pep
112	7	2.1	504	6	ABP72340	ABP72340 Human myo	185	6	1.8	22	2	AAW54884	AAW54884 Ion chann
113	7	2.1	504	6	ABO44238	ABO44238 Human TIG	186	6	1.8	22	2	AAW66527	AAW66527 Amphiphil
114	7	2.1	504	7	ADES4517	ADES4517 Human PRO	187	6	1.8	23	6	ABO12236	ABO12236 Human zin
115	7	2.1	508	3	AAW32190	AAW32190 Arabidops	188	6	1.8	23	6	ABO11924	ABO11924 Human zin
116	7	2.1	514	4	AAW90586	AAW90586 Human sec	189	6	1.8	24	7	ADB47956	ADB47956 Novel hum
117	7	2.1	514	5	ABG65472	ABG65472 Human alb	190	6	1.8	28	4	ABB16060	ABB16060 Human ner
118	7	2.1	523	7	ADC01554	ADC01554 Enterohae	191	6	1.8	28	5	ABG76643	ABG76643 Human SOU
119	7	2.1	535	4	ABBS8106	ABBS8106 Drosophil	192	6	1.8	29	5	AAW15122	AAW15122 Corticotr
120	7	2.1	550	4	ABG67171	ABG67171 Novel hum	193	6	1.8	33	2	AAW43220	AAW43220 Peptide #
121	7	2.1	551	4	AAW85251	AAW85251 Plant thi	194	6	1.8	34	4	AAW37060	AAW37060 Peptide #
122	7	2.1	551	4	AAW85252	AAW85252 Plant thi	195	6	1.8	34	4	AAW76952	AAW76952 Human bon
123	7	2.1	557	3	AAW32076	AAW32076 Arabidops	196	6	1.8	34	4	AAW64127	AAW64127 Human bra
124	7	2.1	567	3	AAW32075	AAW32075 Arabidops	197	6	1.8	34	4	ABG58614	ABG58614 Human liv
125	7	2.1	576	3	AAW32074	AAW32074 Arabidops	198	6	1.8	34	5	ABG46059	ABG46059 Human pep
126	7	2.1	614	3	AAW39771	AAW39771 Arabidops	199	6	1.8	38	2	AAW63517	AAW63517 Bovine ne
127	7	2.1	622	3	AAW32189	AAW32189 Arabidops	200	6	1.8	39	5	ADB67852	ADB67852 Human lun
128	7	2.1	647	3	AAW53000	AAW53000 Human sec	201	6	1.8	39	7	AAW63516	AAW63516 Bovine ne
129	7	2.1	658	6	AAO23969	AAO23969 Human opt	202	6	1.8	40	2	AAW63518	AAW63518 Bovine ne
130	7	2.1	685	4	ABW71254	ABW71254 Drosophil	203	6	1.8	40	2	AAW91027	AAW91027 Transplan
131	7	2.1	731	4	AAW74206	AAW74206 Protein e	204	6	1.8	40	5	AAU91026	AAU91026 Transplan
132	7	2.1	731	7	ABR82939	ABR82939 Arabidops	205	6	1.8	41	6	ADA89676	ADA89676 Staphyloc
133	7	2.1	760	4	ABG25929	ABG25929 Novel hum	206	6	1.8	41	6	AAW38318	AAW38318 Peptide #
134	7	2.1	831	5	ABG91571	ABG91571 Purine/py	207	6	1.8	42	4	AAW31757	AAW31757 Peptide #
135	7	2.1	855	5	AAE17313	AAE17313 Human pro	208	6	1.8	42	4	ABG53156	ABG53156 Human liv
136	7	2.1	873	5	AAE24242	AAE24242 Human 577	209	6	1.8	42	4	AAW14123	AAW14123 Peptide #
137	7	2.1	888	5	AAE17312	AAE17312 Human pro	210	6	1.8	43	4	ABW33068	ABW33068 Peptide #
138	7	2.1	891	6	ABU42062	ABU42062 Protein e	211	6	1.8	43	4	AAW26530	AAW26530 Peptide #
139	7	2.1	896	4	ABG22638	ABG22638 Novel hum	212	6	1.8	43	4	ABW27896	ABW27896 Peptide #
140	7	2.1	922	5	AAU76150	AAU76150 Rice lipo	213	6	1.8	43	4	AAW94970	AAW94970 Human rep
141	7	2.1	961	5	ABW09658	ABW09658 Amino aci	214	6	1.8	43	4	ABW18537	ABW18537 Protein #
142	7	2.1	962	2	AAW44194	AAW44194 Rat NMDA	215	6	1.8	43	4	AAW53865	AAW53865 Human bra
143	7	2.1	979	5	ABW09659	ABW09659 Amino aci	216	6	1.8	43	4	AAW53865	AAW53865 Human bra
144	7	2.1	997	5	ABW09657	ABW09657 Amino aci	217	6	1.8	43	4	AAW53865	AAW53865 Human tes
145	7	2.1	1015	5	ABW09660	ABW09660 Amino aci	218	6	1.8	43	4	AAW01861	AAW01861 Peptide #
146	7	2.1	1053	6	ABU41624	ABU41624 Protein e	219	6	1.8	43	4	ABW95680	ABW95680 Human tes
147	7	2.1	1072	4	AAW70871	AAW70871 C albican	220	6	1.8	43	5	ABG35900	ABG35900 Human pep
148	7	2.1	1148	6	ABP58228	ABP58228 Human cel	221	6	1.8	44	3	AAW18732	AAW18732 Zea maye
149	7	2.1	1185	2	AAW23593	AAW23593 Human LYS	222	6	1.8	44	3	AAU14265	AAU14265 Human nov
150	7	2.1	1239	2	AAW45945	AAW45945 Glutamic	223	6	1.8	45	4	AAO04708	AAO04708 Human pol
151	7	2.1	1292	5	ABW77986	ABW77986 Amino aci	224	6	1.8	45	4	ABG09887	ABG09887 Novel hum
152	7	2.1	1537	5	ABG76501	ABG76501 DNA encod	225	6	1.8	46	2	AAW92823	AAW92823 Human tra
153	7	2.1	1545	2	AAW23595	AAW23595 Murine Ly	226	6	1.8	46	2	AAW68326	AAW68326 Human tra
154	7	2.1	1745	4	AAW90551	AAW90551 Human sec	227	6	1.8	46	3	AAW52813	AAW52813 Human dig
155	7	2.1	1745	5	ABG65473	ABG65473 Human alb	228	6	1.8	47	4	AAW92813	AAW92813 Human pol
156	7	2.1	1762	5	AAU99587	AAU99587 Novel hum	229	6	1.8	47	4	AAU22668	AAU22668 Novel hum
157	7	2.1	1762	7	AAO10834	AAO10834 Human cel	230	6	1.8	47	4	AAU22668	AAU22668 Novel hum
158	7	2.1	1766	4	ABW10246	ABW10246 Human cdn	231	6	1.8	47	7	ADB32508	ADB32508 Human nov
159	7	2.1	1766	5	ABP66833	ABP66833 Human pol	232	6	1.8	48	7	ADB78262	ADB78262 Endometri
160	7	2.1	1838	4	ABW64225	ABW64225 Drosophil	233	6	1.8	49	3	AAW34514	AAW34514 Human sec
161	7	2.1	2013	3	AAW18265	AAW18265 Plasmodiu	234	6	1.8	50	4	AAW75342	AAW75342 Human col
162	7	2.1	3064	6	ABO14706	ABO14706 Novel hum	235	6	1.8	50	4	AAU45565	AAU45565 Propionib
163	7	2.1	3433	2	AAW22017	AAW22017 Urothphin.	236	6	1.8	50	4	AAU42299	AAU42299 Propionib
164	7	2.1	3788	2	AAW23594	AAW23594 Murine Ly	237	6	1.8	50	6	ABW42084	ABW42084 Propionib
165	6	1.8	9	6	ABP96902	ABP96902 Human CAS	238	6	1.8	50	6	ABW38818	ABW38818 Propionib
166	6	1.8	9	6	ABU03359	ABU03359 Human exp	239	6	1.8	51	4	AAO04380	AAO04380 Human pol
167	6	1.8	10	2	AAW71167	AAW71167 Peptide u	240	6	1.8	52	3	AAW34389	AAW34389 Human sec
168	6	1.8	11	4	AAW70315	AAW70315 Human Chk	241	6	1.8	53	4	AAU53123	AAU53123 Propionib
169	6	1.8	11	6	ABG74903	ABG74903 Human TFR	242	6	1.8	53	6	ABW49642	ABW49642 Propionib
170	6	1.8	14	4	AAW69193	AAW69193 Human Ace	243	6	1.8	54	4	AAO12979	AAO12979 Human pol
171	6	1.8	15	2	AAW07411	AAW07411 HSV-1 TK	244	6	1.8	54	4	AAU47338	AAU47338 Propionib

245 6 1.8 54 5 ABP07511 Human ORF  
246 6 1.8 54 6 ABM43857  
247 6 1.8 55 3 AAG11542 Arabidops  
248 6 1.8 55 5 ABP01202 Human ORF  
249 6 1.8 55 5 ABP25762 Streptoco  
250 6 1.8 56 4 AAM94423 Human rep  
251 6 1.8 56 4 ABG60253  
252 6 1.8 56 4 AAM06243 Peptide #  
253 6 1.8 56 5 ABG61724 Novel ova  
254 6 1.8 57 4 AAO05144 Human pol  
255 6 1.8 57 5 ABP09201 Human ORF  
256 6 1.8 58 4 AAO13551 Human pol  
257 6 1.8 58 4 AAU60830 Propionib  
258 6 1.8 58 6 ABM57349 Propionib  
259 6 1.8 58 7 ADC27505 Eastern e  
260 6 1.8 60 3 AAY95591 Prostate  
261 6 1.8 60 4 AAB64598 Human sec  
262 6 1.8 60 4 AAU59292 Propionib  
263 6 1.8 60 6 ABM55811 Propionib  
264 6 1.8 61 7 ADC95824 E. faeciu  
265 6 1.8 62 5 ABP09331 Human ORF  
266 6 1.8 63 2 AAY60310 Human end  
267 6 1.8 64 4 AAB41468 Peptide #  
268 6 1.8 64 4 AAM75143 Human bon  
269 6 1.8 64 4 ABG56906 Human liv  
270 6 1.8 64 5 ABP05066 Human ORF  
271 6 1.8 65 4 AAM86439 Human imm  
272 6 1.8 65 6 ABP77808 N. gonorr  
273 6 1.8 66 5 ABP39689 Staphyloc  
274 6 1.8 67 6 ABU01159 S. pneumo  
275 6 1.8 68 4 ABM15018 Human ner  
276 6 1.8 69 4 AAO05131 Human pol  
277 6 1.8 69 4 AAO05132 Human pol  
278 6 1.8 69 4 AAO13185 Human pol  
279 6 1.8 69 6 ABP76942 N. gonorr  
280 6 1.8 70 4 ABG26828 Novel hum  
281 6 1.8 71 4 AAU28073 Novel hum  
282 6 1.8 72 3 AAG56374 Arabidops  
283 6 1.8 72 3 AAG50382 Arabidops  
284 6 1.8 72 4 AAO13435 Human pol  
285 6 1.8 72 6 ABU48499 Protein e  
286 6 1.8 73 2 AAY48341 Human pro  
287 6 1.8 73 4 AAO02840 Human pol  
288 6 1.8 73 4 AAU50544 Propionib  
289 6 1.8 73 5 ABP31526 Human ORF  
290 6 1.8 73 6 ABM47063 Propionib  
291 6 1.8 76 5 ABP32314 Human ORF  
292 6 1.8 77 4 AAO10681 Human pol  
293 6 1.8 77 4 ABG18579 Novel hum  
294 6 1.8 77 5 ABP10463 Human ORF  
295 6 1.8 77 5 ABP63917 Human ORF  
296 6 1.8 78 5 ABU10267 Human lun  
297 6 1.8 79 4 AAU66107 Propionib  
298 6 1.8 79 5 ABU51733 Helicobac  
299 6 1.8 79 6 ABM62626 Propionib  
300 6 1.8 80 3 AAB00168 Brain spe

## ALIGNMENTS

RESULT 1  
AAY94247  
ID AAY94247 standard; protein; 337 AA.

XX AAY94247;

AC  
XX  
DT  
XX  
XX  
XX  
DE

10-AUG-2000 (first entry)

Human calcium binding protein hCBP.

XX Human; calcium binding protein; cancer; inflammation; CBP;

KW reproductiv disorder; autoimmune disorder; developmental disorder;

KW seizure disorder; immune disorder; infection.  
XX Homo sapiens.  
OS WO200029580-A1.  
PN 25-MAY-2000.  
PD 12-NOV-1999; 99WO-US07027.  
PF 13-NOV-1999; 98US-00190965.  
PR (INCY-) INCYTE PHARM INC.  
PA Tang YT, Guegler KJ, Corley NC, Gorgone GA;  
PI WPI; 2000-387793/33.  
XX N-PSDB; AAA27332.  
DR Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
XX diagnosis, prevention and treatment of cancers, immune, developmental or  
XX reproductiv disorders.  
XX Claim 1; Fig 1; 72pp; English.

CC The present sequence is the human calcium binding protein hCBP. It was  
CC obtained by screening a coronary artery smooth muscle cDNA library, from  
CC which five overlapping nucleic acids were isolated, sequenced and  
CC expressed to give the protein. The protein and the gene encoding it are  
CC useful for the diagnosis and treatment of the following types of  
CC disorder: cancers (such as adenocarcinomas), reproductive disorders (such  
CC as infertility, ovulatory defects, endometriosis, disruptions of the  
CC oestrus and menstrual cycles, polycystic ovary syndrome and ovarian  
CC hyperstimulation), autoimmune disorders (such as benign prostatic  
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's  
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary  
CC neuropathies, seizure disorders, immune disorders (such as AIDS,  
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's  
CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,  
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative  
CC colitis), and viral, bacterial, fungal, parasitic, protozoal and  
CC helminthic infections

XX Sequence 337 AA;

Query Match 100.0%; Score 337; DB 3; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKMPLEFSKSHKNPAAEIVKILKDNLAILEKODKKTDKASEEVSYSIQAAMKEILCGTNEKE 60  
Db 1 MKKMPLEFSKSHKNPAAEIVKILKDNLAILEKODKKTDKASEEVSYSIQAAMKEILCGTNEKE 60  
QY 61 PPTFAVAQAQELYSGLLVTLIADQLIDFEGKDVDTQIFNNILRRQIGTSPTVEYIS 120  
Db 61 PPTFAVAQAQELYSGLLVTLIADQLIDFEGKDVDTQIFNNILRRQIGTSPTVEYIS 120  
QY 121 AHPHILFMLLKGYEAPQIALRCGIMLRECIHPEHAKILFNSQRPDPFKYVELSTFDIA 180  
Db 121 AHPHILFMLLKGYEAPQIALRCGIMLRECIHPEHAKILFNSQRPDPFKYVELSTFDIA 180  
QY 181 SDAFATEKDLTRHKVLYADVADFLQNYDTTFEDYEKLOSENVTYKQSLKGLGELILDRH 240  
Db 181 SDAFATEKDLTRHKVLYADVADFLQNYDTTFEDYEKLOSENVTYKQSLKGLGELILDRH 240  
QY 241 NFAIMTKYISKPENLKLMMNLRLDKSPNIQFAHFVKFVFAVSPHKTQIPVILLKNQPK 300  
Db 241 NFAIMTKYISKPENLKLMMNLRLDKSPNIQFAHFVKFVFAVSPHKTQIPVILLKNQPK 300  
QY 301 LIEFLSSFOKERTDDEQFADEKNYILIKQIRDLKKTAP 337  
Db 301 LIEFLSSFOKERTDDEQFADEKNYILIKQIRDLKKTAP 337

RESULT 2  
AAB82090 standard; protein; 337 AA.  
XX ID AAB82090 standard; protein; 337 AA.  
XX AC AAB82090;  
XX DT 26-JUN-2001 (first entry)  
XX DE Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP.  
XX KW Human; cerebroprotective; neuroprotective; vaccine; ANIC-BP.  
XX KW Gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;  
XX KW stroke; acute head trauma; multiple sclerosis; spinal cord injury.  
XX OS Homo sapiens.  
XX PN WO200123552-A1.  
XX PD 05-APR-2001.  
XX PF 18-SEP-2000; 2000WO-EP009132.  
XX PR 24-SEP-1999; 99EP-00118848.  
XX PA (MERE ) MERCK PATENT GMBH.  
XX PI Den Daas I, Duecker K;  
XX DR WPI; 2001-308142/32.  
XX DR N-PSDB; AAF86462.  
XX PT Novel human acute neuronal induced calcium binding polypeptide, and  
XX PT polynucleotides encoding them useful for diagnosing or treating stroke,  
XX PT acute head trauma, multiple sclerosis and spinal cord injury.  
XX PS Claim 1; Page 41-42; 45pp; English.  
XX CC The present sequence is the protein sequence for human Acute Neuronal  
XX CC Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and  
XX CC protein are useful for treating stroke, acute head trauma, multiple  
XX CC sclerosis and spinal cord injury. ANIC-BP coding sequence and protein are  
XX CC also useful as vaccines for inducing an immunological response in a  
XX CC mammal  
XX SQ Sequence 337 AA;  
Query Match 100.0%; Score 337; DB 4; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKMPFLSKSHKPAEIVKILKONLALTEKODKTKDASEVSKSLQAMKEILCGTNEKE 60  
Db 1 MKKMPFLSKSHKPAEIVKILKONLALTEKODKTKDASEVSKSLQAMKEILCGTNEKE 60  
QY 61 PPTAEVAQAQELYSGLLVTLADIQLIDFEGKDVDTQIFNNILRRQIGTRSPVTEYIS 120  
Db 61 PPTAEVAQAQELYSGLLVTLADIQLIDFEGKDVDTQIFNNILRRQIGTRSPVTEYIS 120  
QY 121 AHPHILFMLKGYEAPQIALRCGIMLRECIHPEPLAKIILFNSQFRDFFKVELSTFDIA 180  
Db 121 AHPHILFMLKGYEAPQIALRCGIMLRECIHPEPLAKIILFNSQFRDFFKVELSTFDIA 180  
QY 181 SDAFATPKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKQSLKLLGELLIDRH 240  
Db 181 SDAFATPKDLLTRHKVLVADFLQNYDTIFEDYEKLLQSENYYTKQSLKLLGELLIDRH 240  
QY 241 NFAIMTKYISKPNLKLMMNLLRDKSPNIQFEAFHFVKFVAFSPHKTQPIVEILLKNQPK 300  
Db 241 NFAIMTKYISKPNLKLMMNLLRDKSPNIQFEAFHFVKFVAFSPHKTQPIVEILLKNQPK 300  
QY 301 LIEFLSFKERTDDBOFADEKKNYLIKQIRDLKKTAP 337  
|||

Db 301 LIEFLSFKERTDDBOFADEKKNYLIKQIRDLKKTAP 337  
RESULT 3  
AAM39078  
XX ID AAM39078 standard; protein; 337 AA.  
XX AC AAM39078;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 2223.  
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;  
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX KW leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US034263.  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00498725.  
XX PR 25-APR-2000; 2000US-0052317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00662191.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX PI Zhou P, Goodrich R, Drmanac RT;  
XX DR WPI; 2001-442253/47.  
XX DR N-PSDB; AAI58234.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such  
XX PT as central nervous system injuries.  
XX PS Example 4; SEQ ID NO 2223; 10078pp; English.  
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide  
XX CC of the invention may be used to treat diseases of the peripheral nervous  
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and  
XX CC localised neuropathies and central nervous system diseases, such as  
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX CC utilisation of the activities such as: Immune system suppression,  
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and  
XX CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX CC part of the printed specification  
XX SQ Sequence 337 AA;  
Query Match 100.0%; Score 337; DB 4; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 14408; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent polynucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 289 AA;

Query Match 55.8%; Score 188; DB 4; Length 289;  
Best Local Similarity 99.7%; Pred. No. 1.5e-175;  
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 49 MKEILCGNEKPPEVAQAQAEYSSGLVTLADIQLIDFEGKDVTOIFNNILRRQ 108  
DB 1 MKEILCGNEKPPEVAQAQAEYSSGLVTLADIQLIDFEGKDVTOIFNNILRRQ 60

QY 109 IGTRSPTEVEYISAHPHILFMLLKGVEAPQIALRCGIMLRCEIRHEPLAKIILFSNQERDF 168  
DB 61 IGTRSPTEVEYISAHPHILFMLLKGVEAPQIALRCGIMLRCEIRHEPLAKIILFSNQERDF 120

QY 169 FKYEELSTFDIASDAFATFKOLLTRHKVLDVADFLQNDYDTTFEDYEKLLQSENVTYTKQS 228  
DB 121 FKYEELSTFDIASDAFATFKOLLTRHKVLDVADFLQNDYDTTFEDYEKLLQSENVTYTKQS 180

QY 229 LKLLGELLDRHNFAMTKYISKPENLKMNNLLRDKSPNQFQFAHFVKVFPVSPHKTQ 286  
DB 181 LKLLGELLDRHNFAMTKYISKPENLKMNNLLRDKSPNQFQFAHFVKVFPVSPHKTQ 240

QY 289 PIVEILLKNQPKLIBFLSSFOKERTDDQFADEKNYLIKQIRDLKKTAP 337  
DB 241 PIVEILLKNQPKLIBFLSSFOKERTDDQFADEKNYLIKQIRDLKKTAP 289

RESULT 6  
AA94248  
ID AA94248 standard; protein; 341 AA.  
AC AA94248;  
XX  
XX 10-AUG-2000 (first entry)  
XX Mouse calcium binding protein MO25.  
XX  
XX Mouse; calcium binding protein; cancer; inflammation; MO25; CBP;  
XX reproductive disorder; autoimmune disorder; developmental disorder;  
XX seizure disorder; immune disorder; infection.  
XX Mus sp.

XX WO200029580-A1.  
XX 25-MAY-2000.  
XX 12-NOV-1999; 99WO-US027027.  
XX 13-NOV-1998; 98US-00190965.  
XX (INCY-) INCYTE PHARM INC.  
XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;  
XX WPI; 2000-387793/33.  
XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g. diagnosis, prevention and treatment of cancers, immune, developmental or reproductive disorders.  
XX Disclosure; Page 66-67; 72pp; English.  
XX The present sequence is the mouse calcium binding protein MO25. It was used in a sequence alignment to identify human calcium binding protein hCBP. The hCBP protein and the gene encoding it are useful for the diagnosis and treatment of the following types of disorder: cancers (such as adenocarcinomas), reproductive disorders (such as infertility, ovulatory defects, endometriosis, disruptions of the oestrus and menstrual cycles, polycystic ovary syndrome and ovarian hyperstimulation), autoimmune disorders (such as benign prostatic hyperplasia and prostatitis), developmental disorders (such as Cushing's syndrome, muscular dystrophy and gonadal dysgenesis), hereditary neuropathies, seizure disorders, immune disorders (such as AIDS, allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves' disease, multiple sclerosis, psoriasis, rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative colitis), and viral, bacterial, fungal, parasitic, protozoal and helminthic infections

XX SQ Sequence 341 AA;

Query Match 6.8%; Score 23; DB 3; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266  
DB 245 IMTKYISKPENLKMNNLLRDKS 267

RESULT 7  
AAB48970  
ID AAB48970 standard; protein; 341 AA.  
XX AAB48970;  
XX 27-MAR-2001 (first entry)  
XX Human ANIC-BP (acute neuronal induced calcium-binding protein).  
XX Human; acute neuronal induced calcium-binding protein; ANIC-BP;  
XX MO25 homologue; HymA homologue; drug screening; stroke;  
XX acute head trauma; multiple sclerosis; spinal cord injury; vaccine;  
XX cerebroprotective; neuroprotective.  
XX Homo sapiens.  
XX WO200078947-A1.  
XX 28-DEC-2000.  
XX 14-JUN-2000; 2000WO-EP005457.  
XX 22-JUN-1999; 99EP-00112024.

XX PA (MERE ) MERCK PATENT GMBH.  
 XX PI Den Daas I, Fischer V, Seyfried C, Von Melchner L;  
 XX PS WPI; 2001-102721/11.  
 XX DR N-PSDB; AAC91772.  
 XX PT Novel acute neuronal induced calcium binding protein, useful for treating  
 XX PT acute head trauma, stroke, multiple sclerosis and spinal cord injury.  
 XX PS Claim 2; Page 37; 50pp; English.  
 XX CC The invention relates to human acute neuronal induced calcium-binding  
 CC protein (ANIC-BP) and to nucleic acid encoding it. The invention also  
 CC relates to expression systems and recombinant host cells comprising ANIC-  
 CC BP DNA, the recombinant production of ANIC-BP, antibodies specific for  
 CC ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin Fc  
 CC region, and methods of screening for modulators of ANIC-BP function. ANIC  
 CC -BP has homology and structural similarity to HymA and Mo25 proteins.  
 CC ANIC-BP proteins and nucleotides are useful for treating stroke and acute  
 CC head trauma, multiple sclerosis and spinal cord injury. ANIC-BP proteins  
 CC are useful in screening assays, for identifying membrane bound or soluble  
 CC receptors, and also in vaccines. ANIC-BP nucleotides are useful as  
 CC diagnostic reagents, as tools for tissue expression studies, for  
 CC chromosome localisation studies, as genetic vaccines, and in the  
 CC generation of transgenic animals. The present sequence represents human  
 CC ANIC-BP  
 XX CC  
 XX SQ Sequence 341 AA;  
 Query Match 6.8%; Score 23; DB 4; Length 341;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 244 IMTKYISKPENLKLMMNLLRDKS 266  
 DB 245 IMTKYISKPENLKLMMNLLRDKS 267  
 RESULT 8  
 AAB20387  
 ID AAB20387 standard; protein; 350 AA.  
 XX AC AAB20387;  
 XX DT 11-JUN-2001 (first entry)  
 XX DE Human acute neuronal induced calcium binding protein ANIC-BP-1B.  
 XX KW Acute neuronal induced calcium binding protein; ANIC-BP-1B;  
 KW splice variant; human; stroke; head trauma; Parkinson's disease;  
 KW Alzheimer's disease; multiple sclerosis; spinal cord injury;  
 KW cerebroprotective; antiparkinsonian; neurotropic; neuroprotective; therapy;  
 KW diagnosis; vaccine.  
 XX OS Homo sapiens.  
 XX OS  
 XX PN WO200125423-A1.  
 XX PD 12-APR-2001.  
 XX PF 28-SEP-2000; 2000WO-EP009475.  
 XX PR 04-OCT-1999; 99EP-00119113.  
 XX PA (MERE ) MERCK PATENT GMBH.  
 XX PI Duecker K, Den Daas I;  
 XX PT WPI; 2001-266306/27.  
 XX DR N-PSDB; AAF30688.  
 XX

PT Novel human acute neuronal induced calcium-binding protein like protein  
 PT splice variant, useful for treating stroke, acute head trauma,  
 PT Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal cord  
 PT injury.  
 XX Claim 2; Page 44-45; 49pp; English.  
 XX CC The present sequence is that of a novel human acute neuronal induced  
 CC calcium binding protein-like protein splice variant, ANIC-BP-1B. The  
 CC protein shows homology to other members of the calcium binding protein  
 CC family, including ANIC-BP, a protein discovered by mRNA differential  
 CC display that is upregulated in a rat model of head trauma. ANIC-BP and  
 CC ANIC-BP-1B differ in their C-terminal portions. The variant protein could  
 CC serve as a novel drug target. The invention provides ANIC-BP-1B  
 CC polynucleotides (see AAF30688) and polypeptides, expression vectors, host  
 CC cells and antibodies, as well as methods for producing the protein and  
 CC for treating or preventing disorders associated with expression of the  
 CC protein by inhibiting or activating the action of ANIC-BP-1B. Diseases  
 CC that may be treated include stroke and acute head trauma, Parkinson's  
 CC disease, Alzheimer's disease, multiple sclerosis and spinal cord injury.  
 CC The polynucleotides and polypeptides can also be used in diagnostic  
 CC assays and in vaccines, and to identify agonists and antagonists useful  
 CC for treating conditions associated with ANIC-BP-1B imbalance  
 XX SQ Sequence 350 AA;  
 Query Match 6.8%; Score 23; DB 4; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 244 IMTKYISKPENLKLMMNLLRDKS 266  
 DB 245 IMTKYISKPENLKLMMNLLRDKS 267  
 RESULT 9  
 AAE10858  
 ID AAE10858 standard; protein; 496 AA.  
 XX AC AAE10858;  
 XX DT 18-DEC-2001 (first entry)  
 XX DE Gal4-human ANIC-BP-1 fusion protein.  
 XX KW Human; acute neuronal induced calcium binding protein type 1 ligand;  
 KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;  
 KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;  
 KW gene therapy; fusion protein; Gal4 protein.  
 XX OS Homo sapiens.  
 XX OS Unidentified.  
 XX OS Chimeric.  
 XX PN WO200170771-A2.  
 XX PD 27-SEP-2001.  
 XX PF 20-MAR-2001; 2001WO-EP003149.  
 XX PR 21-MAR-2000; 2000EP-00106110.  
 XX PA (MERE ) MERCK PATENT GMBH.  
 XX PI Den Daas I, Duecker K, Hock B;  
 XX DR WPI; 2001-607519/69.  
 XX PT Novel acute neuronal induced calcium binding protein type 1 ligand  
 PT polypeptides, useful in the treatment of stroke, head trauma, multiple  
 PT sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord  
 PT injury.  
 XX

PS Disclosure; Page 42-44; 46pp; English.

XX The invention relates to human acute neuronal induced calcium binding

CC protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.

CC Sequences of the invention are useful for treating human diseases

CC including stroke, head trauma, multiple sclerosis, Parkinson's disease,

CC Alzheimer's disease and spinal cord injury. They are also useful as

CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound

CC soluble receptors. Polynucleotides of the invention are useful as

CC Alzheimer's disease and spinal cord injury. They are also useful as

CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound

CC soluble receptors. Polynucleotides of the invention are useful as

CC diagnostic reagents, for chromosome localization studies, and as valuable

CC tools for tissue expression studies. They are also useful in gene

CC therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein

CC comprising the Gal4 protein and a C-terminally linked human ANIC-BP-1

CC protein

XX Sequence 496 AA;

SQ Query Match 6.8%; Score 23; DB 4; Length 496;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKLMMNLLRDKS 266

DB 400 IMTKYISKPENLKLMMNLLRDKS 422

RESULT 10

AAE10859

ID AAE10859 standard; protein; 552 AA.

XX AC AAE10859;

DT 18-DEC-2001 (first entry)

XX DE LexA-human ANIC-BP-1 fusion protein.

XX Human; acute neuronal induced calcium binding protein type 1 ligand;

KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;

KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;

KW gene therapy; fusion protein; LexA protein.

XX OS Homo sapiens.

OS Unidentified.

OS Chimeric.

XX PH Key Location/Qualifiers

FT Region 1..202

FT /note= "LexA protein"

FT 203..552

FT /note= "Human ANIC-BP-1 protein"

FT WO200170771-A2.

PN 27-SEP-2001.

XX 20-MAR-2001; 2001WO-EP003149.

XX 21-MAR-2000; 2000EP-00106110.

XX (MERE ) MERCK PATENT GMBH.

XX Den Daas I, Duecker K, Hock B;

XX WPI; 2001-607519/69.

XX Novel acute neuronal induced calcium binding protein type 1 ligand

PT polypeptides, useful in the treatment of stroke, head trauma, multiple

PT sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord

PT injury.

XX Disclosure; Page 44-46; 46pp; English.

XX The invention relates to human acute neuronal induced calcium binding

CC protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.

CC Sequences of the invention are useful for treating human diseases

CC including stroke, head trauma, multiple sclerosis, Parkinson's disease,

CC Alzheimer's disease and spinal cord injury. They are also useful as

CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound

CC soluble receptors. Polynucleotides of the invention are useful as

CC diagnostic reagents, for chromosome localization studies, and as valuable

CC tools for tissue expression studies. They are also useful in gene

CC therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein

CC comprising the LexA protein and a C-terminally linked human ANIC-BP-1

CC protein

XX Sequence 552 AA;

SQ Query Match 6.8%; Score 23; DB 4; Length 552;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKLMMNLLRDKS 266

DB 456 IMTKYISKPENLKLMMNLLRDKS 478

RESULT 11

ABG23844

ID ABG23844 standard; protein; 354 AA.

XX AC ABG23844;

DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #23835.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSB-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS88031.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 54203; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 354 AA;  
  
Query Match 5.0%; Score 17; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 8.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 219 SENYVTKRQSLKLGEL 235  
Db 233 SENYVTKRQSLKLGEL 249  
|||||  
  
RESULT 12  
AAY94249  
ID AAY94249 standard; protein; 339 AA.  
XX  
AC AAY94249;  
XX  
DT 10-AUG-2000 (first entry)  
XX  
DE Drosophila calcium binding protein DMO25.  
XX  
KW Drosophila; calcium binding protein; cancer; inflammation; DMO25; CBP;  
KW reproductive disorder; autoimmune disorder; developmental disorder;  
KW seizure disorder; immune disorder; infection.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200029580-A1.  
XX  
PD 25-MAY-2000.  
XX  
PF 12-NOV-1999; 99WO-US027027.  
XX  
PR 13-NOV-1998; 98US-00190365.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;  
XX WPI; 2000-387793/33.  
XX  
PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
PT diagnosis, prevention and treatment of cancers, immune, developmental or  
PT reproductive disorders.  
XX  
PS Disclosure; Page 67-68; 72pp; English.  
XX  
CC The present sequence is the Drosophila calcium binding protein DMO25. It  
CC was used in a sequence alignment to identify human calcium binding  
CC protein hCBP. The hCBP protein and the gene encoding it are useful for  
CC the diagnosis and treatment of the following types of disorder: cancers  
CC (such as adenocarcinomas), reproductive disorders (such as infertility,  
CC ovulatory defects, endometriosis, disruptions of the oestrus and  
CC menstrual cycles, polycystic ovary syndrome and ovarian  
CC hyperstimulation), autoimmune disorders (such as benign prostatic  
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's  
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary  
CC neuropathies, seizure disorders, immune disorders (such as AIDS,  
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's  
CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,  
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative  
CC colitis), and viral, bacterial, fungal, parasitic, protozoal and  
CC helminthic infections

XX  
SQ Sequence 339 AA;  
  
Query Match 4.5%; Score 15; DB 3; Length 339;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 105 LRRQIGTRSPVVEYI 119  
Db 103 LRRQIGTRSPVVEYI 117  
|||||  
  
RESULT 13  
ABB60392  
ID ABB60392 standard; protein; 339 AA.  
XX  
AC ABB60392;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 7968.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL04495.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 7968; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 339 AA;  
  
Query Match 4.5%; Score 15; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 105 LRRQIGTRSPVVEYI 119  
Db 103 LRRQIGTRSPVVEYI 117  
|||||  
  
RESULT 14  
AAY94250

ID AAY94250 standard; protein; 377 AA.  
 XX AC AAY94250;  
 XX DT 10-AUG-2000 (first entry)  
 XX DE C. elegans yeast-like calcium binding protein.  
 XX KW Calcium binding protein; cancer; inflammation; yeast-like CBP; CBP;  
 XX KW reproductive disorder; autoimmune disorder; developmental disorder;  
 XX KW seizure disorder; immune disorder; infection.  
 XX OS Caenorhabditis elegans.  
 XX PN WO200029580-A1.  
 XX PD 25-MAY-2000.  
 XX PF 12-NOV-1999; 99WO-US027027.  
 XX PR 13-NOV-1998; 98US-00190965.  
 XX PA (INCY-) INCYTE PHARM INC.  
 XX PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;  
 XX WPI; 2000-387793/33.  
 XX PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.  
 XX PT diagnosis, prevention and treatment of cancers, immune, developmental or  
 XX PT reproductive disorders.  
 XX PS Disclosure; Page 68-69; 72pp; English.  
 CC The present sequence is the C. elegans yeast-like CBP. It was used in a  
 CC sequence alignment to identify human calcium binding protein hCBP. The  
 CC hCBP protein and the gene encoding it are useful for the diagnosis and  
 CC treatment of the following types of disorder: cancers (such as  
 CC adenocarcinomas), reproductive disorders (such as infertility, ovulatory  
 CC defects, endometriosis, disruptions of the oestrus and menstrual cycles,  
 CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune  
 CC disorders (such as benign prostatic hyperplasia and prostatitis),  
 CC developmental disorders (such as Cushing's syndrome, muscular dystrophy  
 CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,  
 CC immune disorders (such as AIDS, allergies, anaemia, asthma,  
 CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'  
 CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis, viral,  
 CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,  
 CC bacterial, fungal, parasitic, protozoal and helminthic infections  
 XX SQ Sequence 377 AA;  
 Query Match 4.2%; Score 14; DB 3; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 105 LRRQIGTRSPVEY 118  
 |||||  
 DB 119 LRRQIGTRSPVEY 132  
 RESULT 15  
 ABG23843  
 ID ABG23843 standard; protein; 383 AA.  
 XX AC ABG23843;  
 XX DT 18-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #23834.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX PD N-PSDB; AAS88030.  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 XX PT diagnostics, forensics, gene mapping, identification of mutations  
 XX PT responsible for genetic disorders or other traits and to assess  
 XX PT biodiversity.  
 XX PS Claim 20; SEQ ID NO 54202; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 383 AA;  
 Query Match 3.6%; Score 12; DB 4; Length 383;  
 Best Local Similarity 100.0%; Pred. No. 0.0072;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 63 TEAVAQLAQL 74  
 |||||  
 DB 332 TEAVAQLAQL 343  
 RESULT 16  
 ABG25372  
 ID ABG25372 standard; protein; 639 AA.  
 XX AC ABG25372;  
 XX DT 18-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #25363.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.

```

XX WO200175067-A2.
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS89559.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 55731; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 639 AA;
XX
Query Match 3.6%; Score 12; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 TEAVAQLAQLY 74
DB 203 TEAVAQLAQLY 214
|||||
|||||

RESULT 17
AAG41153
ID AAG41153 standard; protein; 139 AA.
XX
XX AAG41153;
AC
XX
DT 18-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 51164.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX
XX

```

```

PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 04-MAY-1999; 99US-0132048P.
XX 05-MAY-1999; 99US-0132407P.
XX 06-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 03-JUN-1999; 99US-0137222P.
XX 04-JUN-1999; 99US-0137528P.
XX 07-JUN-1999; 99US-0137502P.
XX 08-JUN-1999; 99US-0137724P.
XX 10-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 14-JUN-1999; 99US-0138847P.
XX 16-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 17-JUN-1999; 99US-0139453P.
XX 18-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 21-JUN-1999; 99US-0139763P.
XX 22-JUN-1999; 99US-0139817P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.

```



```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142977P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155488P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.4%; Score 8; DB 3; Length 139;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 MLRECIRH 152
DB 1 MLRECIRH 8

RESULT 18
AAG41152
ID AAG41152 standard; protein; 148 AA.
XX AC AAG41152;
XX 18-OCT-2000 (first entry)
DT Zea mays protein fragment SEQ ID NO: 51163.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX Zea mays subsp. mays.
XX EF1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
```

PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132048P.  
PR 05-MAY-1999; 99US-0132407P.  
PR 06-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.

```
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.4%; Score 8; DB 3; Length 148;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 145 MLRECIRH 152
Db 10 MLRECIRH 17
|||||

RESULT 19
AAG41151
ID AAG41151 standard; protein; 154 AA.
XX AC AAG41151;
XX DT 18-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 51162.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Zea mays subsp. mays.
XX PN EF1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PF 25-FEB-1999; 99US-0121825P.
XX PF 05-MAR-1999; 99US-0123180P.
XX PF 09-MAR-1999; 99US-0123548P.
XX PF 23-MAR-1999; 99US-0125798P.
XX PF 25-MAR-1999; 99US-0126264P.
XX PF 29-MAR-1999; 99US-0126785P.
XX PF 01-APR-1999; 99US-0127462P.
XX PF 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
```



CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 464 AA;

Query Match 2.4%; Score 8; DB 4; Length 464;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 ADQLIDF 91  
Db 166 ADQLIDF 173

RESULT 21  
ABBS58067  
ID ABB58067 standard; protein; 923 AA.  
XX AC ABB58067;  
XX DT 26-MAR-2002 (first entry)  
XX DE *Drosophila melanogaster* polypeptide SEQ ID NO 993.  
XX KW *Drosophila*; developmental biology; cell signalling; insecticide;  
XX OS pharmaceutical.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX DR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WI; 2001-656860/75.  
XX N-PSDB; ABL02170.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from *Drosophila* and for elucidating cell signaling and cell-cell  
XX interactions.  
XX Disclosure; SEQ ID NO 993; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 923 AA;

Query Match 2.4%; Score 8; DB 4; Length 923;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLVTL 82  
Db 449 SSGLLVTL 456

RESULT 22  
ABG03535  
ID ABG03535 standard; protein; 44 AA.  
XX AC ABG03535;  
XX DT 13-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #3526.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS *Homo sapiens*.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WI; 2001-639362/73.  
XX N-PSDB; AAS67722.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX Claim 20; SEQ ID NO 33894; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have application in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABO0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 44 AA;

Query Match 2.1%; Score 7; DB 4; Length 44;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159337P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84  
 Db 2 LLVTLIA 8

RESULT 24

ABB53042  
 ID ABB53042 standard; protein; 51 AA.  
 AC ABB53042;  
 DT 11-FEB-2002 (first entry)  
 DE Escherichia coli polypeptide SEQ ID NO 1453.

KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;  
 immunosuppressive; extra-intestinal infection; phylogeny; meningitis;  
 KW septicemia; non-diarrhoeal infection; septicaemia;  
 pyelonephritis; antibiotic resistance.

OS Escherichia coli.

PN WO200166572-A2.

XX 13-SEP-2001.

XX 12-MAR-2001; 2001WO-EP003445.

XX 10-MAR-2000; 2000FR-00003145.

PR 02-FEB-2001; 2001FR-00001449.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;

DR WPI; 2001-550253/61.

PT A library of DNA fragments of Escherichia coli strains for the phylogenetic  
 determination of a given strain comprises polynucleotides of nature B2/D+  
 A-.

PS Example 6; Fig 6; 646pp; English.

The invention relates to a library of DNA fragments of Escherichia coli  
 strains comprising polynucleotides (ABA98577-ABA88729 and ABA89533) and  
 encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature  
 B2/D+A-. The polynucleotides have potential antiinflammatory,  
 antibacterial and immunosuppressive activity as part of pharmaceutical  
 compositions used to treat, palliate or prevent extra-intestinal E. coli  
 infections. The polypeptides are useful for determining the phylogenetic  
 group of a given E. coli strain. These polypeptides can detect and treat  
 an undesired development of E. coli, particularly an extra-intestinal  
 infection that include systemic and non-diarrhoeal infections such as  
 septicemia, pyelonephritis and meningitis this is particularly  
 advantageous as bacterial resistance is increasing with the more frequent  
 use of broad spectrum antibiotics

XX Sequence 51 AA;

Query Match 2.1%; Score 7; DB 4; Length 51;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKQ 298

Db 13 EILLKQ 19

RESULT 25

ABP05605

ID ABP05605 standard; protein; 83 AA.

AC ABP05605;

DT 24-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:11192.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;





RESULT 27  
AAM86795  
ID AAM86795 standard; protein; 98 AA.  
XX  
AC AAM86795;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:14388.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytotactic; gene therapy; vaccine; metastasis.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI  
 XX WPI; 2001-483426/52.  
 DR N-PSDB; AAK59576.  
 XX  
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 XX Claim 11; SEQ ID NO 14388; 3071pp + Sequence Listing; English.  
 PS  
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK82170 to AAK51921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 XX Sequence 98 AA;  
 SQ  
 Query Match 2.1%; Score 7; DB 4; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 150 IRHEPLA 156  
 Db |||||||  
 8 IRHEPLA 14  
 RESULT 28  
 AAG89310  
 ID AAG89310 standard; protein; 99 AA.  
 XX  
 AC AAG89310;  
 XX  
 DT 11-SEP-2001 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 430.  
 XX  
 XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;  
 KW GENSET.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200142451-A2.  
 PN  
 XX 14-JUN-2001.  
 PD  
 XX

PF 07-DEC-2000; 2000WO-IB001938.  
 XX  
 PR 08-DEC-1999; 99US-0169629P.  
 PR 06-MAR-2000; 2000US-0187470P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;  
 XX WPI; 2001-367870/38.  
 DR N-PSDB; AAK64913.  
 XX  
 XX Full length GENSET human nucleic acids encoding potentially secreted  
 PT proteins, useful in gene therapy and vaccination against a variety of  
 PT diseases, and for diagnosis of those diseases.  
 XX  
 PS Claim 21; Page 898; 921pp; English.  
 XX  
 CC The invention relates to full length GENSET human nucleic acids encoding  
 CC potentially secreted proteins. The nucleic acids and the polypeptides  
 CC they encode may be used in the prevention, treatment and diagnosis of  
 CC diseases associated with inappropriate GENSET gene expression. For  
 CC example, they be used to treat disorders associated with decreased GENSET  
 CC gene expression by rectifying mutations or deletions in a patient's  
 CC genome that affect the activity of GENSET or by supplementing the  
 CC patient's own production of GENSET polypeptides. Conversely, antisense  
 CC nucleic acid molecules may be administered to down regulate GENSET  
 CC expression by binding with the cells' own genes and preventing their  
 CC expression. The sense and antisense nucleic acids may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar nucleic acid sequences in samples, and hence to determine which  
 CC patients may be in need of restorative therapy. The GENSET polypeptides  
 CC may be used as antigens in the production of antibodies and in assays to  
 CC identify modulators (agonists and antagonists) of GENSET polypeptide  
 CC expression and activity. The present sequence is a GENSET polypeptide of  
 CC the invention  
 XX  
 SQ Sequence 99 AA;  
 Query Match 2.1%; Score 7; DB 4; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 59 KEPTEA 65  
 Db |||||||  
 87 KEPTEA 93  
 RESULT 29  
 ABP76204  
 ID ABP76204 standard; protein; 99 AA.  
 XX  
 AC ABP76204;  
 XX  
 DT 21-FEB-2003 (first entry)  
 XX  
 DE Human GENSET protein SEQ ID 530.  
 XX  
 KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;  
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;  
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;  
 KW central nervous system; cardiovascular; gastrointestinal.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283898-A1.  
 XX  
 XX 24-OCT-2002.  
 PD  
 XX 18-APR-2001; 2001WO-IB000914.  
 PF  
 XX 18-APR-2001; 2001WO-IB000914.  
 PR  
 XX

PA (GEST ) GENSET.  
XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;  
XX WPI; 2003-075548/07.  
DR New GENSET polynucleotides and polypeptides, useful for treating heavy  
XX metal toxicity, cancer, inflammatory diseases, immune disorders, and the  
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the  
PT toxicity.  
XX Claim 14; Page 565; 735pp; English.  
PS The present invention relates to novel GENSET polynucleotides (ABZ36404-  
XX ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides  
CC and polypeptides are useful in screening and diagnostic assays for  
CC abnormal GENSET expression and/or biological activity. They are also  
CC useful for screening of compounds for treating or preventing GENSET-  
CC related disorders, such as heavy metal toxicity, cancer, inflammatory  
CC diseases, immune disorders, and the neuromuscular, central nervous system  
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity  
XX  
XX Sequence 99 AA;  
XX  
XX Query Match 2.1%; Score 7; DB 6; Length 99;  
XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 59 KEPTEA 65  
XX 87 KEPTEA 93  
XX  
XX RESULT 30  
XX ABP76061  
XX ID ABP76061 standard; protein; 99 AA.  
XX AC ABP76061;  
XX DT 21-FEB-2003 (first entry)  
XX DE Human GENSET protein SEQ ID 268.  
XX  
XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;  
XX gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;  
XX inflammatory disease; immune disorder; neuromuscular; toxicity;  
XX central nervous system; cardiovascular; gastrointestinal.  
XX  
XX Homo sapiens.  
XX  
XX WO200283898-A1.  
XX PD 24-OCT-2002.  
XX PF 18-APR-2001; 2001WO-IB000914.  
XX  
XX 18-APR-2001; 2001WO-IB000914.  
XX (GEST ) GENSET.  
XX  
XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;  
XX WPI; 2003-075548/07.  
XX  
XX New GENSET polynucleotides and polypeptides, useful for treating heavy  
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the  
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the  
PT toxicity.  
XX  
XX Claim 14; Page 444; 735pp; English.  
XX  
XX The present invention relates to novel GENSET polynucleotides (ABZ36404-  
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides

CC and polypeptides are useful in screening and diagnostic assays for  
CC abnormal GENSET expression and/or biological activity. They are also  
CC useful for screening of compounds for treating or preventing GENSET-  
CC related disorders, such as heavy metal toxicity, cancer, inflammatory  
CC diseases, immune disorders, and the neuromuscular, central nervous system  
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity  
XX  
XX Sequence 99 AA;  
XX  
XX Query Match 2.1%; Score 7; DB 6; Length 99;  
XX Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 59 KEPTEA 65  
XX 87 KEPTEA 93  
XX  
XX RESULT 31  
XX AA009956  
XX ID AA009956 standard; protein; 107 AA.  
XX AC AA009956;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human polypeptide SEQ ID NO 23848.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorders; arthritis; inflammation.  
XX  
XX Homo sapiens.  
XX  
XX WO200164835-A2.  
XX PD 07-SEP-2001.  
XX PF 26-FEB-2001; 2001WO-US004927.  
XX  
XX 28-FEB-2000; 2000US-00515126.  
XX 18-MAY-2000; 2000US-00577409.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
XX N-PSDB; AAI89887.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
XX Claim 20; SEQ ID NO 23848; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: the sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 107 AA;  
XX  
XX Query Match 2.1%; Score 7; DB 4; Length 107;

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVPK 278  
Db 36 EAFHVPK 42

## RESULT 32

AG74439  
ID AAG74439 standard; protein; 113 AA.

XX AC AAG74439;  
XX DT 03-SEP-2001 (first entry)  
XX DE Human colon cancer antigen protein SEQ ID NO:5203.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
XX KW colorectal carcinoma.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US026524.

XX PR 29-SEP-1999; 99US-0157137P.

XX PR 03-NOV-1999; 99US-0163280P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX DR WPI; 2001-235357/24.

XX DR N-PSDB; AAH33870.

XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX PS Claim 11; Page 6889; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patient's own production of P. Additionally, N may be used  
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922

XX SQ Sequence 113 AA;

Query Match 2.1%; Score 7; DB 4; Length 113;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 BEVSKSL 46

Db 51 BEVSKSL 57

## RESULT 33

ABW71944  
ID ABW71944 standard; protein; 115 AA.

XX AC ABW71944;

XX DT 20-NOV-2003 (first entry)

XX DE Staphylococcus aureus protein #1184.

XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;  
XX KW enzymatic assay; antibiotic target.

XX OS Staphylococcus aureus.

XX PN WO200294868-A2.

XX PD 28-NOV-2002.

XX PF 27-MAR-2002; 2002WO-IB002637.

XX PR 27-MAR-2001; 2001GB-00007661.

XX PA (CHIR-) CHIRON SPA.

XX PI Masignani V, Mora M, Scarselli M;

XX DR WPI; 2003-120786/11.

XX DR N-PSDB; ACF73504.

XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or  
XX PT preventing Staphylococcal infection, specifically an infection caused by  
XX PT S. aureus, e.g. sepsis.

XX PS Claim 1; SEQ ID NO 2368; 49pp; English.

CC The invention relates to novel genes and encoded proteins from  
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a  
CC nucleic acid encoding the protein, or an antibody to the protein, is  
CC useful as a pharmaceutical, particularly as a vaccine for treating or  
CC preventing infection due to Staphylococcus bacteria, specifically an  
CC infection caused by S. aureus. The composition is particularly useful for  
CC treating or preventing sepsis in a patient. The composition can also be  
CC used for diagnostics. The protein is also used in an assay for enzymatic  
CC studies and as a target for antibiotics. This sequence represents one of  
CC the novel S. aureus proteins of the invention

XX SQ Sequence 115 AA;

Query Match 2.1%; Score 7; DB 6; Length 115;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KDNLAAIL 28

Db 78 KDNLAAIL 84

## RESULT 34

ABB65844

ID ABB65844 standard; protein; 139 AA.

XX AC ABB65844;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 24324.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.

XX OS Drosophila melanogaster.

```
XX WO200171042-A2.
XX
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX
XX PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL09947.
XX
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX
XX PS Disclosure; SEQ ID NO 24324; 21pp + Sequence Listing; English.
XX
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 139 AA;

Query Match          2.1%; Score 7; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ISKPENL 255
DB 11 ISKPENL 17
|||||

RESULT 35
ABU31422
ID ABU31422 standard; protein; 141 AA.
XX
XX AC ABU31422;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #16949.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Klebsiella pneumoniae.
XX
XX PN WO200277183-A2.
XX
XX PD 03-OCT-2002.
XX
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX
XX PR 06-SEP-2001; 2001US-00948993.
XX
XX PR 25-OCT-2001; 2001US-0342923P.
XX
XX PR 08-FEB-2002; 2002US-00072851.
XX
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX N-PSDB; ACA35292.
XX
XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX PS Claim 25; SEQ ID NO 59346; 1766pp; English.
XX
XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 141 AA;

Query Match          2.1%; Score 7; DB 6; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKIL 21
DB 107 AEIVKIL 113
|||||

RESULT 36
AAV29337
ID AAY29337 standard; protein; 176 AA.
XX
XX AC AAY29337;
XX
XX DT 29-SEP-1999 (first entry)
XX
XX DE Human secreted protein clone g5984_13 alternate reading frame protein.
XX
XX KW Human; secreted protein; nutrition; cytokine; cell proliferation;
XX KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
XX KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
XX KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
XX KW tumour invasion suppressor; tumour inhibition.
XX
XX OS Homo sapiens.
XX
```

```

PN WO9937674-A1.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1999; 99WO-US001404.
XX
PR 22-JAN-1998; 98US-0072134P.
XX
PR 20-JAN-1999; 99US-00235609.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Merberg D;
PI Treacy M, Agostino MJ, Steininger RJ, Wong GG, Clark HF, Fechtel K;
XX WPI; 1999-458682/38.
XX N-PSDB; AAX90444.
XX
PT New polynucleotides encoding secreted human proteins derived from, e.g.
PT fetal brain potentially used as immunostimulators.
XX
PS Disclosure; Page 134; 139pp; English.
XX
CC The present sequence represents a human secreted protein. Human secreted
CC protein polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful for
CC gene therapy
XX
SQ Sequence 176 AA;
Query Match 2.1%; Score 7; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 231 LLGELIL 237
Db 98 LLGELIL 104
|||||
RESULT 37
AAU39088
ID AAU39088 standard; protein; 176 AA.
XX
AC AAU39088;
XX
XX
DT 16-JAN-2002 (first entry)
XX
DE Human secreted protein gg894_1 #2.
XX
XX Human; secreted protein; antiinflammatory; immunosuppressive; nootropic;
XX neuroprotective; antiarthritic; antimicrobial; vulnery; cytostatic;
XX antidiabetic; virucide; antiinfertility; anticonvulsant; vasotropic;
XX antiparkinsonian; immunostimulant; dermatological; antirheumatic;
XX antitumor; antiulcer; osteopathic; tranquiliser; cerebroprotective;
XX cytokine; cell proliferation; cell differentiation; immune deficiency;
XX severe combined immunodeficiency; SCID; tumour; autoimmune disorder;
XX multiple sclerosis; rheumatoid arthritis; SCID; tumour; autoimmune disease;
XX myeloid deficiency; wound healing; ulcer; periodontal disease;
XX osteoporosis; osteoarthritis; Alzheimer's disease; Parkinson's disease;
XX Huntington's disease; infection; cardiac disease; stroke; sepsis;
XX inflammatory bowel disease; contraceptive; immunogen; food supplement;
XX vaccine.
XX
XX Homo sapiens.
XX

```

---

```

PN WO200175068-A2.
XX
PD 11-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US009369.
XX
PR 30-MAR-2000; 2000US-00539330.
XX
PR 04-DEC-2000; 2000US-00729674.
XX
PA (GEMY ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie E, Collins-Racie LA, Evans C;
PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG, Clark H;
PI Fechtel K, Merberg D;
XX WPI; 2001-639363/73.
XX N-PSDB; AASS9272.
XX
PT Secreted human proteins, useful as vaccine for treating various diseases
PT such as autoimmune disorders (e.g. multiple sclerosis), and nervous
PT system disorders (e.g. stroke).
XX
PS Disclosure; Page 617-618; 619pp; English.
XX
CC The invention relates to novel human secreted proteins, the nucleic acids
CC encoding them. The protein may exhibit cytokine, cell proliferation or
CC cell differentiation activity or may induce production of other cytokines
CC in certain cell populations and may exhibit immune stimulating or immune
CC suppressing activity, which is useful for the treatment of various immune
CC deficiencies and disorders e.g. severe combined immunodeficiency (SCID),
CC autoimmune disorders e.g. multiple sclerosis, systemic lupus
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation.
CC The proteins are also useful in the treatment of diseases and disorders
CC including tissue, skin and organ transplantation and in graft-versus-host
CC diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid
CC cell deficiencies, wound healing and tissue repair, in the treatment of
CC burns, incisions and ulcers; as well as in treatment of periodontal
CC disease, osteoporosis or osteoarthritis, mediated by inflammatory
CC processes, diseases of the peripheral nervous system, Alzheimer's,
CC Parkinson's disease, Huntington's disease, amyotrophic lateral
CC sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and
CC central nervous system vessel e.g. stroke, sepsis, inflammatory bowel
CC disease, ulcers, bone regeneration. The protein, having activin- or
CC inhibin-related activities is useful as a contraceptive based on the
CC ability of inhibins to decrease fertility in female mammals and decrease
CC spermatogenesis in male mammals. The proteins and nucleic acids are also
CC useful as food supplements. The present sequence represents a secreted
CC protein of the invention
XX
SQ Sequence 176 AA;
Query Match 2.1%; Score 7; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 231 LLGELIL 237
Db 98 LLGELIL 104
|||||
RESULT 38
ABB55797
ID ABB55797 standard; protein; 176 AA.
XX
AC ABB55797;
XX
XX
DT 14-FEB-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 280.
XX
XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
XX immune disorder; bacterial infection; fungal infection; cancer; tumour;
XX autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;

```



osteoporosis; osteoarthritis; nervous system disorder; neuropathy;  
 Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;  
 haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;  
 ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;  
 Crohn's disease; cytosatic; anti-inflammatory; immunomodulator;  
 neuroprotective; haemostatic; thrombolytic; anti-inflammatory.

OS Homo sapiens.

XX US2001039335-A1.

XX 08-NOV-2001.

PF 04-DEC-2000; 2000US-00729674.

XX 26-NOV-1997; 97US-0126425P.

PR 04-DEC-1997; 97US-0067454P.

PR 20-DEC-1997; 97US-0068379P.

PR 02-JAN-1998; 98US-0070346P.

PR 07-JAN-1998; 98US-0070643P.

PR 08-JAN-1998; 98US-0070755P.

PR 13-JAN-1998; 98US-0071304P.

PR 22-JAN-1998; 98US-0072134P.

PR 30-JAN-1998; 98US-0073095P.

PR 18-FEB-1998; 98US-0075038P.

PR 23-NOV-1998; 98US-00197886.

PR 30-MAR-2000; 2000US-00539330.

XX (JACO/) JACOBS K.

XX (MCCO/) MCCOY J M.

XX (LAVA/) LAVALLIE E R.

XX (COLL/) COLLINS-RACIE L A.

XX (EVAN/) EVANS C.

XX (MERB/) MERBERG D.

XX (TREA/) TREACY M.

XX (AGOS/) AGOSTINO M J.

XX (STEI/) STEININGER R J.

XX (SPAU/) SPAULDING V.

XX (WONG/) WONG G G.

XX (CLAR/) CLARK H.

XX (FECH/) FECHTEL K.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;

XX Wong GG, Clark H, Fechtel K;

XX WPI; 2002-040725/05.

XX New secreted proteins and encoding polynucleotides, useful in gene

XX therapies, particularly for preventing or treating autoimmune disorders,

XX cancer, graft-versus-host disease, wound, osteoporosis, stroke or

XX inflammations.

XX Disclosure; Page 343-344; 349pp; English.

XX The invention relates to isolated polynucleotides (ABA90876-ABA90968 and

XX ABA90980) and encoded proteins (ABB55698-ABB55800), especially

XX polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and

XX proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in

XX clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1

XX are deposited with the American Type Culture Collection (ATCC) with

XX accession number 98599. The polynucleotides and encoded polypeptides have

XX cytosatic, anti-inflammatory, immunomodulator, vulnery,

XX neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic

XX and anti-inflammatory activity and acting as cytokine modulators,

XX haematopoiesis regulatory tissue growth modulators and/or cadherin

XX suppressors. The polypeptides and polynucleotides are useful in gene

XX therapies, particularly for preventing, treating or ameliorating any of

XX the following diseases: immune deficiency and disorders; e.g. bacterial

XX or fungal infections, autoimmune disorders, cancer, systemic lupus

XX erythematosus or graft-versus-host disease; myeloid or lymphoid cell

XX deficiencies; wound, burns, incisions and ulcers, osteoporosis or

XX osteoarthritis; central and peripheral nervous system diseases and

CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's  
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;  
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis  
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,  
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's  
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus  
 XX  
 SQ Sequence 176 AA;

Query Match 2.1%; Score 7; DB 5; Length 176;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELIL 237

Db 98 LLGELIL 104

RESULT 39

AAU31106

ID AAU31106 standard; protein; 192 AA.

XX AAU31106;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #1597.

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

XX vaccination, testing and therapy.

XX Claim 20; Page 403-404; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention

XX Sequence 192 AA;

Query Match 2.1%; Score 7; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 ADLQID 90  
DQ 167 ADLQID 173

## RESULT 40

AAAG05091  
ID AAAG05091 standard; protein; 197 AA.

XX AC AAAG05091;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 1364.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132484P.

XX PR 06-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132487P.

XX PR 07-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

XX PR 27-MAY-1999; 99US-0136392P.

XX PR 28-MAY-1999; 99US-0136782P.

XX PR 01-JUN-1999; 99US-0137222P.

XX PR 03-JUN-1999; 99US-0137528P.

XX PR 04-JUN-1999; 99US-0137502P.

XX PR 07-JUN-1999; 99US-0137724P.

XX PR 08-JUN-1999; 99US-0138094P.

XX PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144844P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145513P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147018P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148311P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149369P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149425P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149923P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160845P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161466P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161932P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Db 42 TFKDLT 48  
 RESULT 41  
 ADC31499  
 ID ADC31499 standard; protein; 199 AA.  
 XX ADC31499;  
 AC ADC31499;  
 XX 18-DEC-2003 (first entry)  
 DT Human novel polypeptide sequence, SEQ ID NO:1581.  
 XX  
 DE Human; diagnostic; drug screening; forensics; gene mapping;  
 XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
 KW ulcers; osteoporosis; autoimmune disease; cancer;  
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;  
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
 KW gene therapy; chromosome 5.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W02003029271-A2.  
 PN  
 XX 10-APR-2003.  
 PD  
 XX  
 XX 24-SEP-2002; 2002WO-US030474.  
 PF  
 XX 24-SEP-2001; 2001US-0324631P.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
 PI Haley-Vicente D, Drmanac RT;  
 XX  
 XX WPI; 2003-371981/35.  
 DR N-PSDB; ADC30528.  
 DR  
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
 PT treating conditions such as neurodegenerative diseases, anemias, platelet  
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 PT cancer.  
 XX  
 PS Claim 20; SEQ ID NO 1581; 1185pp; English.  
 XX  
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
 CC invention also relates to nucleic acid sequences over 99% identical with  
 CC the novel human cDNAs. The invention additionally encompasses expression  
 CC vectors and host cells comprising a nucleic acid of the invention; the  
 CC recombinant production of a polypeptide of the invention; an antibody  
 CC against a polypeptide of the invention; a method of detecting of  
 CC polynucleotides or polypeptides of the invention; and methods of  
 CC identifying a compound which binds to a polypeptide of the invention. The  
 CC invention further discloses methods of preventing, treating or  
 CC ameliorating a medical condition; kits comprising polynucleotide probes  
 CC and/or monoclonal antibodies for carrying out the methods of the  
 CC invention; methods for the identification of compounds that modulate the  
 CC expression or activity of the polynucleotide and/or polypeptide; and 767  
 CC contig sequences corresponding to the cDNA sequences of the invention  
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
 CC -ADC33394). The nucleic acids and polypeptides of the invention are  
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
 CC identification of mutations responsible for genetic disorders or other  
 CC traits, for assessing biodiversity, and in producing many other types of  
 CC data and products dependent on DNA and amino acid sequences. They are  
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
 CC disease and other neurodegenerative diseases, anaemia, platelet  
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 CC cancer. The nucleic acids may also be used as hybridisation probes or

Query Match 2.1%; Score 7; DB 3; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 TFKDLT 192  
 |||||

CC primers, and in the recombinant production of a protein. The polypeptides  
 CC are also useful in generating antibodies, as molecular weight markers,  
 CC and as food supplements. The present sequence represents a specifically  
 CC claimed human polypeptide sequence of the invention. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIDO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 199 AA;

Query Match 2.1%; Score 7; DB 7; Length 199;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGLL 79

DB 57 LYSSGLL 63

RESULT 42

AAW11481  
 ID AAW11481 standard; protein; 205 AA.

XX AC AAW11481;

DT 22-APR-1997 (first entry)

DE D. immitis mature venom allergen antigen 5-like protein PDiVA205.

XX Venom allergen antigen 5-like gene; VA5; helminth; parasite; nDiVA615;

KW PDiVA205; vaccine; heartworm.

XX OS Dirofilaria immitis.

XX PN WO9637218-A1.

XX PD 28-NOV-1996.

XX PF 23-MAY-1996; 96WO-US007709.

XX PR 23-MAY-1995; 95US-00450944.

XX PA (HESK-) HESKA CORP.

XX PI Tripp CA, Wisniewski N;

XX DR WPI; 1997-020935/02.

XX DR N-P8DB; AAT51379.

PT Nucleic acid encoding helminth venom allergen antigen 5-like protein -  
 PT Pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp. in  
 PT vaccines to prevent helminth infection.

XX Claim 19; Page 91-92; 124pp; English.

XX Dirofilaria immitis mature venom allergen antigen 5-like protein VA5  
 CC (AAW11481), or PDiVA205, is capable of eliciting an immune response  
 CC (cellular and/or humoral) in animals. It is encoded by a cDNA molecule  
 CC nDiVA615 (AAT51379). VA5-like proteins (see also AAW11480-86) can be used  
 CC in vaccines to prevent helminth infection

SQ Sequence 205 AA;

Query Match 2.1%; Score 7; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 RDLKXTA 336

DB 93 RDLKXTA 99

RESULT 43

AAG30715

ID AAG30715 standard; protein; 213 AA.

XX AC AAG30715;

DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36771.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132484P.

XX PR 06-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

XX PR 27-MAY-1999; 99US-0136392P.

XX PR 28-MAY-1999; 99US-0136578P.

XX PR 01-JUN-1999; 99US-0137222P.

XX PR 03-JUN-1999; 99US-0137528P.

XX PR 04-JUN-1999; 99US-0137502P.

XX PR 07-JUN-1999; 99US-0137724P.

XX PR 08-JUN-1999; 99US-0138094P.

XX PR 10-JUN-1999; 99US-0138540P.

XX PR 10-JUN-1999; 99US-0138847P.

XX PR 14-JUN-1999; 99US-0139119P.

XX PR 16-JUN-1999; 99US-0139452P.

XX PR 16-JUN-1999; 99US-0139453P.

XX PR 17-JUN-1999; 99US-0139492P.

XX PR 18-JUN-1999; 99US-0139454P.

XX PR 18-JUN-1999; 99US-0139455P.

XX PR 18-JUN-1999; 99US-0139456P.

XX PR 18-JUN-1999; 99US-0139457P.

XX PR 18-JUN-1999; 99US-0139458P.

XX PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139761P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0145386P.  
 PR 02-AUG-1999; 99US-0145388P.  
 PR 02-AUG-1999; 99US-0145389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 06-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148313P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.

PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151086P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 22-OCT-1999; 99US-0160800P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVEK 278  
 |||||  
 Db 148 EAFHVEK 154

RESULT 44  
 AAG23888  
 ID AAG23888 standard; protein; 213 AA.  
 XX  
 AC AAG23888;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27359.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX EPI033405-A2.  
XX 06-SEP-2000.  
PD  
PD  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134376P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.





```
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
Query Match 2.1%; Score 7; DB 3; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 272 EAFHVPK 278
Db 148 EAFHVPK 154
RESULT 46
AAU25691
ID AAU25691 standard; protein; 215 AA.
XX
AC AAU25691;
XX
DT 18-DEC-2001 (first entry)
XX
DE G protein-coupled receptor, nGPCR-2105.
XX
KW Human; mental disorder; thyroid disease; renal failure; anorexia;
KW inflammatory condition; Crohn's disease; rheumatoid arthritis; HIV;
KW autoimmune disorder; schizophrenia; migraine; stroke; dementia; obesity;
KW depression; Parkinson's disease; Alzheimer's disease; viral infection;
KW Huntington's disease; human immunodeficiency virus; type 2 diabetes;
KW anorexia; hypotension; hypertension; thrombosis; myocardial infarction;
KW atherosclerosis; cancer; sexual dysfunction; G protein-coupled receptor;
nGPCR.
XX
OS Homo sapiens.
XX
PN WO200162924-A2.
```

```

XX 30-AUG-2001.
PD
XX
XX
XX 23-FEB-2001; 2001WO-US005989.
PF
XX
XX
XX 24-FEB-2000; 2000US-0184602P.
PR
XX
XX 24-FEB-2000; 2000US-0184604P.
PR
XX
XX 24-FEB-2000; 2000US-0184606P.
PR
XX
XX 24-FEB-2000; 2000US-0184689P.
PR
XX
XX 24-FEB-2000; 2000US-0184690P.
PR
XX
XX 24-FEB-2000; 2000US-0184710P.
PR
XX
XX 24-FEB-2000; 2000US-0184712P.
PR
XX
XX 24-FEB-2000; 2000US-0184715P.
PR
XX
XX 24-FEB-2000; 2000US-0184716P.
PR
XX
XX 24-FEB-2000; 2000US-0184725P.
PR
XX
XX 24-FEB-2000; 2000US-0184822P.
PR
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
PA
XX
XX Vogeli G, Wood LS, Parodi LA, Lind P;
PI
XX
XX WPI; 2001-570632/54.
DR
XX
XX N-PSDB; AAS43001.
DR
XX
XX Novel nucleic acid and encoded nGPCR-x, used to screen for compounds for
PT
XX
XX use in the treatment of mental disorders, such as Alzheimer's disease, or
PT
XX
XX Parkinson's disease.
PT
XX
XX Claim 31; Page 82; 263pp; English.
PS
XX
XX The invention relates to novel isolated human G protein-coupled receptors
CC
XX
XX (nGPCR-x). The nGPCR-x can be used for screening compounds which can be
CC
XX
XX used to treat mental disorders, thyroid disease, renal failure,
CC
XX
XX inflammatory conditions such as Crohn's disease, rheumatoid arthritis,
CC
XX
XX autoimmune disorders, schizophrenia, migraine, stroke, dementia,
CC
XX
XX depression, Parkinson's disease, Alzheimer's disease, and Huntington's
CC
XX
XX disease. They may also be used for treating viral infections such as
CC
XX
XX human immunodeficiency virus (HIV), type 2 diabetes, obesity, anorexia,
CC
XX
XX hypotension, hypertension, thrombosis, myocardial infarction,
CC
XX
XX atherosclerosis, cancer, and sexual dysfunction. AAU25617-AAU25726
CC
XX
XX represent the amino acid sequences of novel human G protein-coupled
CC
XX
XX receptors, nGPCR-2031 to nGPCR-2140 respectively, as described in the
CC
XX
XX invention
CC
XX
XX Sequence 215 AA;
SQ
XX
XX
XX Query Match 2.1%; Score 7; DB 4; Length 215;
XX
XX Best Local Similarity 100.0%; Pred. No. 3.4e+02;
XX
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX Qy 69 LAQELYS 75
XX
XX Db 21 LAQELYS 27
XX
XX
XX RESULT 47
XX
XX ID ADA54954 standard; protein; 217 AA.
XX
XX AC ADA54954;
XX
XX AC
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Human protein, SEQ ID 2522.
XX
XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
XX
XX Gene therapy; human; secretory protein; membrane proteins; cancer;
XX
XX Inflammatory disease; osteoporosis; neurological disease.
XX
XX OS Homo sapiens.
XX
XX OS
XX
XX PN EP1293569-A2.
XX
XX
XX 30-AUG-2001.
PD
XX
XX
XX 21-MAR-2003.
PF
XX
XX
XX 21-MAR-2002; 2002EP-00006586.
PR
XX
XX
XX 14-SEP-2001; 2001JP-00328381.
PR
XX
XX
XX 24-JAN-2002; 2002US-0350435P.
PR
XX
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI
XX
XX Yanamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI
XX
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX
XX WPI; 2003-395539/38.
DR
XX
XX N-PSDB; ADA53315.
DR
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT
XX
XX and/or membrane proteins, useful for developing medicines for diseases in
PT
XX
XX which the gene is involved, or as target molecules for gene therapy.
PT
XX
XX Claim 14; SEQ ID NO 2522; 205pp; English.
PS
XX
XX The present invention relates to novel human secretory or membrane
CC
XX
XX proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC
XX
XX ADA54071). The coding sequences are useful in the gene therapy of
CC
XX
XX diseases caused by abnormalities of the proteins, e.g. cancer,
CC
XX
XX inflammatory diseases, osteoporosis or neurological disease.
CC
XX
XX Sequence 217 AA;
SQ
XX
XX
XX Query Match 2.1%; Score 7; DB 6; Length 217;
XX
XX Best Local Similarity 100.0%; Pred. No. 3.4e+02;
XX
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX Qy 24 NLAIIEK 30
XX
XX Db 21 NLAIIEK 27
XX
XX
XX RESULT 48
XX
XX ID AAW11479 standard; protein; 221 AA.
XX
XX AC AAW11479;
XX
XX AC
XX
XX DT 22-APR-1997 (first entry)
XX
XX DE D. immitis venom allergen antigen 5-like protein PDIvA221.
XX
XX
XX Venom allergen antigen 5-like gene; VAS; helminth; parasite; nDiVA833;
XX
XX PDIvA221; vaccine; heartworm.
XX
XX
XX Dirofilaria immitis.
XX
XX
XX Key Location/Qualifiers
XX
XX Peptide 1..15
XX
XX Protein 17..221
XX
XX /label= Sig_peptide
XX
XX /label= Mat_protein
XX
XX /note= "mature protein is represented by PDIvA205"
XX
XX
XX WO9637218-A1.
XX
XX
XX 28-NOV-1996.
XX
XX
XX 23-MAY-1996; 96WO-US007709.
XX
XX
XX 23-MAY-1995; 95US-00450944.
XX
XX
XX (HESK-) HESKA CORP.
XX
XX
XX Tripp CA, Wisniewski N;
XX
XX

```

```

XX WPI; 1997-020935/02.
DR N-PSDB; AAT51376, AAT51378.
XX
XX Nucleic acid encoding helminth venom allergen antigen 5-like protein -
PT pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp. in
PT vaccines to prevent helminth infection.
XX
XX Claim 13; Page 89-90; 124pp; English.
XX
XX Dirofilaria immitis venom allergen antigen 5-like protein VAS (AAW11479),
CC or PDIVA221, is capable of eliciting an immune response (cellular and/or
CC humoral) in animals. It is encoded by a cDNA clone (AAT51376) isolated
CC from a D. immitis cDNA library by screening with immune dog serum. VAS-
CC like proteins (see also AAW11480-86) can be used in vaccines to prevent
CC helminth infection
XX
XX Sequence 221 AA;
SQ
    Query Match          2.1%; Score 7; DB 2; Length 221;
    Best Local Similarity 100.0%; Pred. No. 3.4e+02;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 330 RDLKXKA 336
Db 109 RDLKXKA 115
    |||||
    |||||
RESULT 49
ABP02921
ID ABP02921 standard; protein; 236 AA.
XX
XX AC ABP02921;
XX
XX 24-JUN-2002 (first entry)
XX
XX Human ORFX protein sequence SEQ ID NO: 5824.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX
XX Homo sapiens.
XX
XX WO200192523-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US010836.
XX
XX 30-MAY-2000; 2000US-0206132P.
XX
XX 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach MD;
XX
XX WPI; 2002-106308/14.
XX
XX N-PSDB; ABN18673.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 5824; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

```

```

CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 236 AA;
SQ
    Query Match          2.1%; Score 7; DB 5; Length 236;
    Best Local Similarity 100.0%; Pred. No. 3.7e+02;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 YVTKRQS 228
Db 215 YVTKRQS 221
    |||||
    |||||
RESULT 50
AAE2494
ID AAE24494 standard; protein; 245 AA.
XX
XX AC AAE24494;
XX
XX 04-OCT-2002 (first entry)
XX
XX Human RATL1d6 transmembrane domain.
XX
XX Human; ubiquitin conjugating enzyme; UBC; RATL1d6; immune disorder;
XX regulated in activated T-lymphocyte 1d6; neuronal disorder; cancer;
XX tumour; lymphoproliferative; cancer; adenocarcinoma; leukaemia; myeloma;
XX sarcoma; neurodegenerative; inflammatory; rheumatoid arthritis; asthma;
XX multiple sclerosis; psoriasis; neuronal; Alzheimer's disease; dementia;
XX depression; epilepsy; acquired immuno deficiency syndrome; allergy; AIDS;
XX anaemia; atopic dermatitis; diabetes mellitus; dermatological;
XX myocardial infarction; renal tubular acidosis; gonadal dysgenesis;
XX dysplasia; cataract; cytostatic; neuroprotective; nootropic; anti-HIV;
XX anticonvulsant; antiinflammatory; Cushing's syndrome; cardiac;
XX ophthalmological; transmembrane domain.
XX
XX Homo sapiens.
XX
XX WO200236741-A2.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-US046559.
XX
XX 30-OCT-2000; 2000US-0244688P.
XX
XX 30-JUL-2001; 2001US-0308706P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Bowen MA, Wu Y, Yang W, Finger JN;
XX WPI; 2002-479758/51.
XX

```

PT Novel ubiquitin conjugating enzyme polypeptide isolated from activated  
PT human T cell, for screening modulators useful for treating cancer, immune  
PT disorder, lymphoproliferative disorder, neurodegenerative disorder.

PS Claim 6; Page 61; 169pp; English.

XX  
CC The invention relates to a novel ubiquitin conjugating enzyme (UBC)  
CC homologue, RAU1d6 (regulated in activated T-lymphocytes 1d6) and its  
CC corresponding nucleic acid. The invention also relates to methods for  
CC treating, diagnosing, preventing and screening for disorders related to  
CC the expression of RAU1d6. UBC is useful for screening for candidate  
CC compounds capable of binding to and/or modulating its activity. UBC is  
CC useful for treating an immune or neuronal disorder in a mammal. The  
CC method is useful for treating a cancer or tumour. It is useful for  
CC suppressing the immune response in a subject requiring the suppression.  
CC It is also useful for treating lymphoproliferative disorder, cancer e.g.  
CC adenocarcinoma, leukaemia, myeloma, sarcoma, etc, neurodegenerative  
CC disorder, inflammatory disorders e.g. rheumatoid arthritis, asthma,  
CC multiple sclerosis, psoriasis, etc, neuronal disorders e.g. Alzheimer's  
CC disease, dementia, depression, epilepsy, etc, immune disorder or immune  
CC related disorders such as acquired immuno deficiency syndrome (AIDS),  
CC allergy, anaemia, atopic dermatitis, diabetes mellitus, myocardial  
CC infarction, etc, developmental disorders e.g. Cushing's syndrome, renal  
CC tubular acidosis, gonadal dysgenesis, dysplasia, cataract, etc. The  
CC present sequence is human RAU1d6 transmembrane domain

XX SQ Sequence 245 AA;

Query Match 2.1%; Score 7; DB 5; Length 245;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
Db 138 NLAILEK 144  
|||||

RESULT 51  
AAU63772  
ID AAU63772 standard; protein; 257 AA.

AC AAU63772;

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #24668.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59635.

PT Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 24967; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 257 AA;

Query Match 2.1%; Score 7; DB 4; Length 257;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGT 111  
Db 107 LRRQIGT 113  
|||||

RESULT 52

ABM60291

ID ABM60291 standard; protein; 257 AA.

AC ABM60291;

DT 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #24967.

XX Acne vulgaris; antisporrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI BarCh B, Vallieve-Douglass J;

XX WPI; 2003-381789/36.

XX N-PSDB; ACP64564.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.

Example 1; SEQ ID NO 24967; 1481pp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 257 AA;

Query Match 2.1%; Score 7; DB 6; Length 257;

Best Local Similarity 100.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LRRQIGT 111

Db 107 LRRQIGT 113

RESULT 53

AA041773

ID AAB41773 standard; protein; 263 AA.

AC AAB41773;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF1537 polypeptide sequence SEQ ID NO:3074.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antiproliferative; antiparkinsonian; nontropic; neuroprotective; anticonvulsant; osteopathic; antarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antihypertensive; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

OS Homo sapiens.

PN WO200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US008621.

XX

PR 31-MAR-1999; 99US-0127607P.

PR 02-APR-1999; 99US-0127636P.

PR 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

PI WPI; 2000-602362/57.

DR N-PSDB; AAC75982.

XX Novel nucleic acids and peptides derived from open reading frame X,

PT neurodegenerative disorders and cardiovascular disease.

PT Claim 11; Page 2294; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX

CC sequences have activities such as: cytostatic; hepatotropic; vulnary;

CC antiproliferative; antiparkinsonian; nontropic; neuroprotective; osteopathic;

CC anticonvulsant; antarthritic; immunosuppressant; immunostimulant;

CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;

CC dermatological; immunosuppressive; antiinflammatory; antibacterial;

CC antiviral; antifungal; antirheumatic; antihypertensive; antianaemic. The

CC sequences can be used for determining the presence of or predisposition

CC to, or preventing or treating pathological conditions associated with an

CC ORFX-associated disorder. The nucleic acids can be used to express ORFX

CC proteins in gene therapy vectors. The proteins and nucleic acids may be

CC used to treat cancers, proliferative disorders, neurodegenerative

CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,

CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

CC storage, systemic lupus erythematosus, severe combined immunodeficiency

CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune

CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and

CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to

CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 263 AA;

Query Match 2.1%; Score 7; DB 3; Length 263;

Best Local Similarity 100.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 24 NLAILEX 30

Db 189 NLAILEX 195

RESULT 54

ABU05677

ID ABU05677 standard; protein; 275 AA.

AC ABU05677;

XX M. tuberculosis and M. leprae marker protein #328.

DT 08-APR-2003 (first entry)

DE Mycobacteriosis; survival; virulence; protective antigen; vaccine;

XX Mycobacterial disease; tuberculosis; leprosy.

KW Mycobacterium tuberculosis.

XX Mycobacterium leprae.

OS WO200274903-A2.

PN 26-SEP-2002.

PD 22-FEB-2002; 2002WO-IB001973.

PF



PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0123788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132048P.  
PR 08-MAY-1999; 99US-0132407P.  
PR 14-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133456P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134321P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 16-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138054P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151067P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.



PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-01592293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 22-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 28-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161931P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVFX 278  
 DB 225 EAFHVFX 231  
 |||||

RESULT 57  
 AAY48312  
 ID AAY48312 standard; protein; 291 AA.  
 XX  
 AC AAY48312;

XX 08-DEC-1999 (first entry)  
 XX Human prostate cancer-associated protein 9.  
 DE  
 DE Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;  
 KW cancer; tissue specificity; human.  
 KW  
 XX Homo sapiens.  
 OS  
 XX DE19811194-AL.  
 PN  
 XX 16-SEP-1999.  
 PD  
 XX 10-MAR-1998; 98DE-01011194.  
 PF  
 XX 10-MAR-1998; 98DE-01011194.  
 XX  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 FA  
 XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
 PI WPI; 1999-519629/44.  
 XX  
 DR N-PSDB; AAZ33482.  
 DR  
 XX

PT New nucleic acid expressed at high level in normal prostatic tissue and  
 PT encoded polypeptides, used to treat cancer and screen for therapeutic  
 PT agents.  
 XX  
 XX Claim 22; 128; 194pp; German.  
 XX  
 CC This invention describes novel nucleic acid sequences (A) that are  
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)  
 CC encoded by (A) are used: (a) for identifying agents for treatment of  
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally where  
 CC expressed by gene therapy methods. (A) is also used to isolate full-  
 CC length genes (for gene therapy) and for recombinant production of (I),  
 CC which can be used to raise specific antibodies. (A) are identified by  
 CC assembly of ESTs (expressed sequence tags) before these are analyzed for  
 CC expression pattern (tissue specificity). This approach eliminates many of  
 CC the false results, as regards tissue specificity, associated with known  
 CC methods that use single (usually short) ESTs. AAY48304-Y48456 represent  
 CC peptides encoded by the expressed sequence tags described in the method  
 CC of the invention  
 XX  
 SQ Sequence 291 AA;

Query Match 2.1%; Score 7; DB 2; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
 DB 95 NLAILEK 101  
 |||||

RESULT 58  
 AAG45274  
 ID AAG45274 standard; protein; 300 AA.

XX  
 AC AAG45274;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 56917.  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX EP1033405-A2.  
 PN  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.



```

PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-01611404P.
PR 25-OCT-1999; 99US-01611405P.
PR 25-OCT-1999; 99US-01611406P.
PR 25-OCT-1999; 99US-0161159P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-01621142P.

Query Match      2.1%; Score 7; DB 3; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVK 278
Db 235 EAFHVK 241

RESULT 59
AAG30714
ID AAG30714 standard; protein; 300 AA.
XX
AC AAG30714;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36770.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

```

```
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154179P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158469P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      2.1%; Score 7; DB 3; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVFK 278
Db 235 EAFHVFK 241

RESULT 60
AAG23886
ID AAG23886 standard; protein; 300 AA.
AC AAG23886;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27357.
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0128785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
```

PR 28-MAY-1999; 99US-0136782P;  
PR 01-JUN-1999; 99US-0137222P;  
PR 03-JUN-1999; 99US-0137528P;  
PR 04-JUN-1999; 99US-0137502P;  
PR 07-JUN-1999; 99US-0137724P;  
PR 08-JUN-1999; 99US-0138094P;  
PR 10-JUN-1999; 99US-0138540P;  
PR 10-JUN-1999; 99US-0138847P;  
PR 14-JUN-1999; 99US-0139119P;  
PR 16-JUN-1999; 99US-0139452P;  
PR 16-JUN-1999; 99US-0139453P;  
PR 17-JUN-1999; 99US-0139492P;  
PR 18-JUN-1999; 99US-0139454P;  
PR 18-JUN-1999; 99US-0139455P;  
PR 18-JUN-1999; 99US-0139456P;  
PR 18-JUN-1999; 99US-0139457P;  
PR 18-JUN-1999; 99US-0139458P;  
PR 18-JUN-1999; 99US-0139459P;  
PR 18-JUN-1999; 99US-0139460P;  
PR 18-JUN-1999; 99US-0139461P;  
PR 18-JUN-1999; 99US-0139462P;  
PR 18-JUN-1999; 99US-0139463P;  
PR 18-JUN-1999; 99US-0139750P;  
PR 18-JUN-1999; 99US-0139763P;  
PR 21-JUN-1999; 99US-0139817P;  
PR 22-JUN-1999; 99US-0139899P;  
PR 23-JUN-1999; 99US-0140353P;  
PR 23-JUN-1999; 99US-0140354P;  
PR 24-JUN-1999; 99US-0140695P;  
PR 28-JUN-1999; 99US-0140823P;  
PR 29-JUN-1999; 99US-0140991P;  
PR 30-JUN-1999; 99US-0141287P;  
PR 01-JUL-1999; 99US-0141842P;  
PR 01-JUL-1999; 99US-0142154P;  
PR 02-JUL-1999; 99US-0142055P;  
PR 06-JUL-1999; 99US-0142390P;  
PR 08-JUL-1999; 99US-0142803P;  
PR 09-JUL-1999; 99US-0142920P;  
PR 12-JUL-1999; 99US-0142977P;  
PR 13-JUL-1999; 99US-0143542P;  
PR 14-JUL-1999; 99US-0143624P;  
PR 15-JUL-1999; 99US-0144005P;  
PR 16-JUL-1999; 99US-0144085P;  
PR 16-JUL-1999; 99US-0144086P;  
PR 19-JUL-1999; 99US-0144325P;  
PR 19-JUL-1999; 99US-0144331P;  
PR 19-JUL-1999; 99US-0144332P;  
PR 19-JUL-1999; 99US-0144333P;  
PR 19-JUL-1999; 99US-0144334P;  
PR 19-JUL-1999; 99US-0144335P;  
PR 20-JUL-1999; 99US-0144352P;  
PR 20-JUL-1999; 99US-0144632P;  
PR 20-JUL-1999; 99US-0144884P;  
PR 21-JUL-1999; 99US-0144814P;  
PR 21-JUL-1999; 99US-0145086P;  
PR 21-JUL-1999; 99US-0145088P;  
PR 22-JUL-1999; 99US-0145087P;  
PR 22-JUL-1999; 99US-0145089P;  
PR 22-JUL-1999; 99US-0145192P;  
PR 22-JUL-1999; 99US-0145145P;  
PR 23-JUL-1999; 99US-0145218P;  
PR 23-JUL-1999; 99US-0145224P;  
PR 26-JUL-1999; 99US-0145276P;  
PR 27-JUL-1999; 99US-0145913P;  
PR 27-JUL-1999; 99US-0145918P;  
PR 27-JUL-1999; 99US-0145919P;  
PR 28-JUL-1999; 99US-0145951P;  
PR 02-AUG-1999; 99US-0146386P;  
PR 02-AUG-1999; 99US-0146388P;  
PR 02-AUG-1999; 99US-0146389P;  
PR 03-AUG-1999; 99US-0147038P;  
PR 04-AUG-1999; 99US-0147204P;  
PR 04-AUG-1999; 99US-0147302P;  
PR 05-AUG-1999; 99US-0147192P;  
PR 05-AUG-1999; 99US-0147260P;  
PR 06-AUG-1999; 99US-0147303P;  
PR 06-AUG-1999; 99US-0147416P;  
PR 09-AUG-1999; 99US-0147493P;  
PR 09-AUG-1999; 99US-0147935P;  
PR 10-AUG-1999; 99US-0148171P;  
PR 11-AUG-1999; 99US-0148319P;  
PR 12-AUG-1999; 99US-0148341P;  
PR 13-AUG-1999; 99US-0148565P;  
PR 13-AUG-1999; 99US-0148684P;  
PR 16-AUG-1999; 99US-0149368P;  
PR 17-AUG-1999; 99US-0149175P;  
PR 18-AUG-1999; 99US-0149426P;  
PR 20-AUG-1999; 99US-0149722P;  
PR 20-AUG-1999; 99US-0149723P;  
PR 20-AUG-1999; 99US-0149929P;  
PR 23-AUG-1999; 99US-0149902P;  
PR 23-AUG-1999; 99US-0149930P;  
PR 25-AUG-1999; 99US-0150566P;  
PR 26-AUG-1999; 99US-0150884P;  
PR 27-AUG-1999; 99US-0151065P;  
PR 27-AUG-1999; 99US-0151066P;  
PR 27-AUG-1999; 99US-0151080P;  
PR 30-AUG-1999; 99US-0151303P;  
PR 31-AUG-1999; 99US-0151438P;  
PR 01-SEP-1999; 99US-0151930P;  
PR 07-SEP-1999; 99US-0152363P;  
PR 10-SEP-1999; 99US-0153070P;  
PR 13-SEP-1999; 99US-0153758P;  
PR 15-SEP-1999; 99US-0154018P;  
PR 16-SEP-1999; 99US-0154039P;  
PR 20-SEP-1999; 99US-0154779P;  
PR 22-SEP-1999; 99US-0155139P;  
PR 23-SEP-1999; 99US-0155486P;  
PR 24-SEP-1999; 99US-0155659P;  
PR 28-SEP-1999; 99US-0156458P;  
PR 29-SEP-1999; 99US-0156596P;  
PR 04-OCT-1999; 99US-0157117P;  
PR 05-OCT-1999; 99US-0157753P;  
PR 06-OCT-1999; 99US-0157865P;  
PR 07-OCT-1999; 99US-0158029P;  
PR 08-OCT-1999; 99US-0158232P;  
PR 12-OCT-1999; 99US-0158369P;  
PR 13-OCT-1999; 99US-0159293P;  
PR 13-OCT-1999; 99US-0159294P;  
PR 13-OCT-1999; 99US-0159295P;  
PR 14-OCT-1999; 99US-0159329P;  
PR 14-OCT-1999; 99US-0159330P;  
PR 14-OCT-1999; 99US-0159331P;  
PR 14-OCT-1999; 99US-0159637P;  
PR 14-OCT-1999; 99US-0159638P;  
PR 18-OCT-1999; 99US-0159584P;  
PR 21-OCT-1999; 99US-0160741P;  
PR 21-OCT-1999; 99US-0160767P;  
PR 21-OCT-1999; 99US-0160768P;  
PR 21-OCT-1999; 99US-0160770P;  
PR 21-OCT-1999; 99US-0160814P;  
PR 21-OCT-1999; 99US-0160815P;  
PR 22-OCT-1999; 99US-0160980P;  
PR 22-OCT-1999; 99US-0160981P;  
PR 22-OCT-1999; 99US-0160989P;  
PR 25-OCT-1999; 99US-0161404P;  
PR 25-OCT-1999; 99US-0161405P;  
PR 25-OCT-1999; 99US-0161406P;  
PR 26-OCT-1999; 99US-0161359P;  
PR 26-OCT-1999; 99US-0161360P;  
PR 26-OCT-1999; 99US-0161361P;  
PR 28-OCT-1999; 99US-0161920P;  
PR 28-OCT-1999; 99US-0161922P;  
PR 28-OCT-1999; 99US-0161993P;  
PR 29-OCT-1999; 99US-0162142P;

Query Match 2.1%; Score 7; DB 3; Length 300;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVFK 278  
|||||  
Db 235 EAFHVFK 241

## RESULT 61

ABM71583  
ID ABM71583 standard; protein; 302 AA.

XX AC ABM71583;

XX DT 20-NOV-2003 (first entry)

XX DE Staphylococcus aureus protein #823.

XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;

XX KW enzymatic assay; antibiotic target.

XX OS Staphylococcus aureus.

XX PN WO200294868-A2.

XX PD 28-NOV-2002.

XX PF 27-MAR-2002; 2002WO-IB002637.

XX PR 27-MAR-2001; 2001GB-00007661.

XX PA (CHIR-) CHIRON SPA.

XX PI Masignani V, Mora M, Scarselli M;

XX DR WPI: 2003-120786/11.

XX DR N-PSDB; ACF73143.

XX PT New Staphylococcus aureus protein, useful as a vaccine for treating or  
preventing Staphylococcal infection, specifically an infection caused by  
S. aureus, e.g. sepsis.

XX PS Claim 1; SEQ ID NO 1646; 49pp; English.

XX CC The invention relates to novel genes and encoded proteins from  
Staphylococcus aureus. A composition comprising the S. aureus protein, a  
nucleic acid encoding the protein, or an antibody to the protein, is  
useful as a pharmaceutical, particularly as a vaccine for treating or  
preventing infection due to Staphylococcus bacteria, specifically an  
infection caused by S. aureus. The composition is particularly useful for  
treating or preventing sepsis in a patient. The composition can also be  
used for diagnostics. The protein is also used in an assay for enzymatic  
studies and as a target for antibiotics. This sequence represents one of  
the novel S. aureus proteins of the invention

XX SQ Sequence 302 AA;

Query Match 2.1%; Score 7; DB 6; Length 302;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 KLIQSEN 221  
|||||  
Db 66 KLIQSEN 72

## RESULT 62

AAG30713  
ID AAG30713 standard; protein; 305 AA.

XX AC AAG30713;

XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36769.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
XX PR 05-MAR-1999; 99US-0123180P.  
XX PR 09-MAR-1999; 99US-0133548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 25-MAR-1999; 99US-0126264P.  
XX PR 29-MAR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 06-APR-1999; 99US-0128234P.  
XX PR 08-APR-1999; 99US-0128714P.  
XX PR 16-APR-1999; 99US-0129845P.  
XX PR 19-APR-1999; 99US-0130077P.  
XX PR 21-APR-1999; 99US-0130449P.  
XX PR 23-APR-1999; 99US-0130510P.  
XX PR 28-APR-1999; 99US-0130891P.  
XX PR 30-APR-1999; 99US-0131449P.  
XX PR 30-APR-1999; 99US-0132048P.  
XX PR 30-APR-1999; 99US-0132407P.  
XX PR 04-MAY-1999; 99US-0132484P.  
XX PR 05-MAY-1999; 99US-0132485P.  
XX PR 06-MAY-1999; 99US-0132486P.  
XX PR 06-MAY-1999; 99US-0132487P.  
XX PR 07-MAY-1999; 99US-0132863P.  
XX PR 11-MAY-1999; 99US-0134256P.  
XX PR 14-MAY-1999; 99US-0134218P.  
XX PR 14-MAY-1999; 99US-0134219P.  
XX PR 14-MAY-1999; 99US-0134221P.  
XX PR 14-MAY-1999; 99US-0134370P.  
XX PR 18-MAY-1999; 99US-0134768P.  
XX PR 19-MAY-1999; 99US-0134941P.  
XX PR 20-MAY-1999; 99US-0135124P.  
XX PR 21-MAY-1999; 99US-0135353P.  
XX PR 24-MAY-1999; 99US-0135629P.  
XX PR 25-MAY-1999; 99US-0136021P.  
XX PR 27-MAY-1999; 99US-0136392P.  
XX PR 28-MAY-1999; 99US-0136782P.  
XX PR 01-JUN-1999; 99US-0137222P.  
XX PR 03-JUN-1999; 99US-0137528P.  
XX PR 04-JUN-1999; 99US-0137502P.  
XX PR 07-JUN-1999; 99US-0137724P.  
XX PR 08-JUN-1999; 99US-0138094P.  
XX PR 10-JUN-1999; 99US-0138540P.  
XX PR 10-JUN-1999; 99US-0138847P.  
XX PR 14-JUN-1999; 99US-0139119P.  
XX PR 16-JUN-1999; 99US-0139452P.  
XX PR 16-JUN-1999; 99US-0139453P.  
XX PR 17-JUN-1999; 99US-0139492P.  
XX PR 18-JUN-1999; 99US-0139454P.  
XX PR 18-JUN-1999; 99US-0139455P.  
XX PR 18-JUN-1999; 99US-0139456P.  
XX PR 18-JUN-1999; 99US-0139457P.  
XX PR 18-JUN-1999; 99US-0139458P.  
XX PR 18-JUN-1999; 99US-0139459P.  
XX PR 18-JUN-1999; 99US-0139460P.  
XX PR 18-JUN-1999; 99US-0139461P.  
XX PR 18-JUN-1999; 99US-0139462P.  
XX PR 18-JUN-1999; 99US-0139463P.

```
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 25-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144386P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147955P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149356P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 305;
Best Local Similarity 100.0%; Pred. NO. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVEK 278
Db 240 EAFHVEK 246

RESULT 63
AAV73387
ID AAY73387 standard; protein; 308 AA.
XX AAY73387;
XX 24-FEB-2000 (first entry)
DT HTRM clone 3340290 protein sequence.
DE HTRM; human transcriptional regulatory molecule; arteriosclerosis; AIDS;
XX arteriosclerosis; cirrhosis; cancer; leukaemia; diabetes mellitus;
KW
```



KW Addison's disease; multiple sclerosis; rheumatoid arthritis; infection;  
 KW trauma; myasthenia gravis; adenocarcinoma; immune disorder; treatment.  
 OS Homo sapiens.  
 XX WO9957144-A2.  
 XX PN 11-NOV-1999.  
 XX 04-MAY-1999; 99WO-US009935.  
 XX 05-MAY-1998; 98US-0084254P.  
 PR 07-AUG-1998; 98US-0095827P.  
 PR 02-OCT-1998; 98US-0102745P.  
 XX (INCY-) INCYTE PHARM INC.  
 XX Hillman JL, Bandman O, Lal P, Yue H, Reddy R, Tang YT;  
 PI Gerstin EH, Patterson C, Baughn MR, Azimzai Y, Lu DAM;  
 XX WPI: 2000-052941/04.  
 DR N-PSDB; AA252472.  
 XX New peptides useful for diagnosis, prevention and treatment of cancer and  
 PT immune disorders.  
 XX Claim 1; Page 144-145; 193pp; English.  
 XX AA73325-Y73389 are human transcriptional regulator molecule (HTRM)  
 CC protein sequences. The HTRM protein and nucleotide sequences are useful  
 CC for preventing or treating disorders associated with decreased expression  
 CC or activity of HTRM which include cell proliferative disorders such as  
 CC arteriosclerosis and cirrhosis; cancers including adenocarcinoma and  
 CC leukaemia; immune disorders such as AIDS, Addison's disease, diabetes  
 CC mellitus, rheumatoid arthritis, multiple sclerosis, systemic lupus  
 CC erythematosus, and myasthenia gravis; infections and trauma. Antagonists  
 CC of the HTRM polypeptides are useful for treating or preventing disorders  
 CC associated with increased expression or activity of HTRMs. HTRM  
 CC polypeptides, their immunogenic fragments or oligopeptides are useful for  
 CC screening libraries of compounds in drug screening techniques.  
 CC Polynucleotides encoding HTRM are useful for blocking the transcription  
 CC of mRNA and regulating gene function by modulating the activity of HTRM.  
 CC Vectors expressing HTRM or agonists can also be used to prevent or treat  
 CC disorder associated with decreased HTRM expression. Antibodies which  
 CC specifically bind HTRM and polynucleotides encoding HTRM are useful for  
 CC diagnosing disorders associated with the expression of HTRM, particularly  
 CC in assays that detect the expression of HTRM. Nucleotide sequences  
 CC encoding HTRM may be useful to generate hybridization probes useful in  
 CC mapping the naturally occurring genomic sequence and to detect  
 CC differences in gene sequences among normal, carrier and affected  
 CC individuals. Using diagnostic assays, cancer can be detected prior to the  
 CC appearance of clinical symptoms and thereby progression of cancer can be  
 CC prevented by aggressive treatment or preventive measures  
 XX Sequence 308 AA;  
 SQ Query Match 2.1%; Score 7; DB 3; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 24 NLALILEK 30  
 Db 144 NLALILEK 150  
 RESULT 64  
 AAG05090  
 ID AAG05090 standard; protein; 320 AA.  
 XX  
 AC AAG05090;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 1363.  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 PN EP1033405-A2.  
 XX 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-00301439.  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 23-MAR-1999; 99US-0123548P.  
 PR 25-MAR-1999; 99US-0125788P.  
 PR 29-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132484P.  
 PR 06-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132863P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140253P.
PR 24-JUN-1999; 99US-0140354P.
PR 25-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 27-JUN-1999; 99US-0140991P.
PR 28-JUN-1999; 99US-0141287P.
PR 29-JUN-1999; 99US-0141842P.
PR 30-JUN-1999; 99US-0142154P.
PR 01-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142390P.
PR 03-JUL-1999; 99US-0142803P.
PR 04-JUL-1999; 99US-0142920P.
PR 05-JUL-1999; 99US-0142977P.
PR 06-JUL-1999; 99US-0143542P.
PR 07-JUL-1999; 99US-0143624P.
PR 08-JUL-1999; 99US-0144005P.
PR 09-JUL-1999; 99US-0144085P.
PR 10-JUL-1999; 99US-0144086P.
PR 11-JUL-1999; 99US-0144325P.
PR 12-JUL-1999; 99US-0144331P.
PR 13-JUL-1999; 99US-0144332P.
PR 14-JUL-1999; 99US-0144333P.
PR 15-JUL-1999; 99US-0144334P.
PR 16-JUL-1999; 99US-0144335P.
PR 17-JUL-1999; 99US-0144352P.
PR 18-JUL-1999; 99US-0144384P.
PR 19-JUL-1999; 99US-0144814P.
PR 20-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 23-JUL-1999; 99US-0145089P.
PR 24-JUL-1999; 99US-0145192P.
PR 25-JUL-1999; 99US-0145145P.
PR 26-JUL-1999; 99US-0145218P.
PR 27-JUL-1999; 99US-0145224P.
PR 28-JUL-1999; 99US-0145276P.
PR 29-JUL-1999; 99US-0145913P.
PR 30-JUL-1999; 99US-0145918P.
PR 31-JUL-1999; 99US-0145919P.
PR 01-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 03-AUG-1999; 99US-0146388P.
PR 04-AUG-1999; 99US-0146389P.
PR 05-AUG-1999; 99US-0147038P.
PR 06-AUG-1999; 99US-0147204P.
PR 07-AUG-1999; 99US-0147302P.
PR 08-AUG-1999; 99US-0147192P.
PR 09-AUG-1999; 99US-0147260P.
PR 10-AUG-1999; 99US-0147303P.
PR 11-AUG-1999; 99US-0147416P.
PR 12-AUG-1999; 99US-0147493P.
PR 13-AUG-1999; 99US-0147935P.
PR 14-AUG-1999; 99US-0148171P.
PR 15-AUG-1999; 99US-0148319P.
PR 16-AUG-1999; 99US-0148341P.
PR 17-AUG-1999; 99US-0148565P.
PR 18-AUG-1999; 99US-0149175P.
PR 19-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 21-AUG-1999; 99US-0149723P.
PR 22-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 24-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 28-AUG-1999; 99US-0151066P.
PR 29-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 02-SEP-1999; 99US-0152363P.
PR 03-SEP-1999; 99US-0153070P.
PR 04-SEP-1999; 99US-0153758P.
PR 05-SEP-1999; 99US-0154018P.
PR 06-SEP-1999; 99US-0154039P.
PR 07-SEP-1999; 99US-0154779P.
PR 08-SEP-1999; 99US-0155139P.
PR 09-SEP-1999; 99US-0155486P.
PR 10-SEP-1999; 99US-0155659P.
PR 11-SEP-1999; 99US-0156458P.
PR 12-SEP-1999; 99US-0156596P.
PR 13-SEP-1999; 99US-0157117P.
PR 14-SEP-1999; 99US-0157753P.
PR 15-SEP-1999; 99US-0157865P.
PR 16-SEP-1999; 99US-0158029P.
PR 17-SEP-1999; 99US-0158232P.
PR 18-SEP-1999; 99US-0158369P.
PR 19-SEP-1999; 99US-0159293P.
PR 20-SEP-1999; 99US-0159294P.
PR 21-SEP-1999; 99US-0159295P.
PR 22-SEP-1999; 99US-0159329P.
PR 23-SEP-1999; 99US-0159330P.
PR 24-SEP-1999; 99US-0159331P.
PR 25-SEP-1999; 99US-0159637P.
PR 26-SEP-1999; 99US-0159638P.
PR 27-SEP-1999; 99US-0159584P.
PR 28-SEP-1999; 99US-0160741P.
PR 29-SEP-1999; 99US-0160767P.
PR 30-SEP-1999; 99US-0160768P.
PR 01-OCT-1999; 99US-0160770P.
PR 02-OCT-1999; 99US-0160814P.
PR 03-OCT-1999; 99US-0160815P.
PR 04-OCT-1999; 99US-0160980P.
PR 05-OCT-1999; 99US-0160981P.
PR 06-OCT-1999; 99US-0160989P.
PR 07-OCT-1999; 99US-0161404P.
PR 08-OCT-1999; 99US-0161405P.
PR 09-OCT-1999; 99US-0161408P.
PR 10-OCT-1999; 99US-0161359P.
PR 11-OCT-1999; 99US-0161360P.
PR 12-OCT-1999; 99US-0161361P.
PR 13-OCT-1999; 99US-0161920P.
PR 14-OCT-1999; 99US-0161922P.
PR 15-OCT-1999; 99US-0161993P.
PR 16-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 320;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 186 TFXDLT 192
DB 165 TFXDLT 171
|||||

RESULT 65
AGS1053
ID AAGS1053 standard; protein; 326 AA.
XX
AC AAGS1053;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 64758.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
```

XX EP1033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123348P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140333P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140655P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.

PR 16-SEP-1999; 99US-01540339P.  
PR 20-SEP-1999; 99US-01547779P.  
PR 22-SEP-1999; 99US-01551339P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158363P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 326;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 EAFHVK 278  
Db 261 EAFHVK 267

RESULT 66  
AAV17865  
ID AAV17865 standard; protein; 332 AA.

XX AC AAV17865;

DT 17-AUG-1999 (first entry)

XX Sulfolobus solfataricus endo-beta-1,4-glucanase Cella.

XX Sulfolobus solfataricus; endoglucanase; Cella; CelB;

XX endo-beta-1,4-glucanase.

XX Sulfolobus solfataricus.

XX Key Location/Qualifiers

XX Peptide 1..23

XX Protein /label= signal

XX /label= Cella

XX DK9900097-A.

XX 12-JAN-1999.  
XX 12-JAN-1999; 99DK-00000097.  
XX 12-JAN-1999; 99DK-00000097.  
XX (NOVO ) NOVO-NORDISK AS.

XX WPI; 1999-279376/24.

XX N-PSDB; AAX80187.

XX Novel endonuclease - produced by Sulfolobus solfataricus.

XX Claim 1; Page 27-28; 36pp; English.

XX The present sequence represents a Sulfolobus solfataricus endoglucanase,  
CC specifically endo-beta-1,4-glucanase, designated Cella. The new endo-beta-  
CC 1,4-glucanase is produced by Sulfolobus solfataricus ATCC 35092

XX Sequence 332 AA;

Query Match 2.1%; Score 7; DB 2; Length 332;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 177 FDIASDA 183

Db 184 FDIASDA 190

RESULT 67

AAV17866

ID AAV17866 standard; protein; 332 AA.

XX AC AAV17866;

DT 17-AUG-1999 (first entry)

XX Sulfolobus solfataricus endo-beta-1,4-glucanase CelB.

XX Sulfolobus solfataricus; endoglucanase; Cella; CelB;

XX endo-beta-1,4-glucanase.

XX Sulfolobus solfataricus.

XX Key Location/Qualifiers

XX Peptide 1..21

XX Protein /label= signal

XX /label= CelB

XX DK9900097-A.

XX 12-JAN-1999.

XX 12-JAN-1999; 99DK-00000097.

XX 12-JAN-1999; 99DK-00000097.

XX (NOVO ) NOVO-NORDISK AS.

XX WPI; 1999-279376/24.

XX N-PSDB; AAX80188.

XX Novel endonuclease - produced by Sulfolobus solfataricus.

XX Claim 1; Page 31-32; 36pp; English.

XX The present sequence represents a Sulfolobus solfataricus endoglucanase,  
CC specifically endo-beta-1,4-glucanase, designated CelB. The new endo-beta-  
CC 1,4-glucanase is produced by Sulfolobus solfataricus ATCC 35092

SQ Sequence 332 AA;

Query Match 2.1%; Score 7; DB 2; Length 332;  
Best Local Similarity 100.0%; Pred. No. 5e+02; 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 FDIASDA 183  
Db 184 FDIASDA 190  
|||||

RESULT 69  
ABG27865  
ID ABG27865 standard; protein; 337 AA.  
XX  
AC ABG27865;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #27856.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00643167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
WPI: 2001-639362/73.  
DR N-PSDB; AAS2052.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 20; SEQ ID NO 58224; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (II) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pdt\_sequences  
XX  
XX Sequence 337 AA;

Query Match 2.1%; Score 7; DB 4; Length 337;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02; 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
Db 31 VAQLAQE 37  
|||||

RESULT 69  
ABB97360  
ID ABB97360 standard; protein; 340 AA.  
XX  
AC ABB97360;  
XX  
DT 27-JUN-2002 (first entry)  
XX  
DE Novel human protein SEQ ID NO: 628.  
XX  
KW Human; antianemic; vulnery; antiinflammatory; immunomodulator;  
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; ESR;  
KW expressed sequence tag.  
XX  
OS Homo sapiens.  
XX  
PN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001WO-US026015.  
XX  
PR 11-SEP-2000; 2000US-00659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
WPI: 2002-292408/33.  
DR N-PSDB; ABN32546.  
XX  
PT An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
PS Example 2; SEQ ID NO 628; 509pp; English.

CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention  
XX  
XX Sequence 340 AA;

Query Match 2.1%; Score 7; DB 5; Length 340;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02; 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
Db 144 NLAILEK 150  
|||||

RESULT 70  
AAG45273  
ID AAG45273 standard; protein; 343 AA.  
XX



```

PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      2.1%; Score 7; DB 3; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.2e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAPHVEK 278
Db 278 EAPHVEK 284
|||||

RESULT 71
ADA32843
ID ADA32843 standard; protein; 343 AA.
XX AC ADA32843;
XX 20-NOV-2003 (first entry)
DT Acinetobacter baumannii protein #4.
DE Acinetobacter baumannii
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW

```

```

KW plant biocontrol agent.
XX Acinetobacter baumannii.
OS US6562958-B1.
PN 13-MAY-2003.
PD 04-JUN-1999; 99US-00328352.
PF 09-JUN-1998; 98US-0088701P.
PR (GENO-) GENOME THERAPEUTICS CORP.
XX Breton G, Bush D;
XX WPI; 2003-576092/54.
DR N-PSDB; ADA28717.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX Example; SEQ ID NO 4130; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 343 AA;

Query Match      2.1%; Score 7; DB 6; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.2e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305
Db 45 PKLIEFL 51
|||||

RESULT 72
ABU23650
ID ABU23650 standard; protein; 344 AA.
XX AC ABU23650;
XX
XX 19-JUN-2003 (first entry)
DT Protein encoded by Prokaryotic essential gene #9177.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Clostridium acetobutylicum.
OS WO200277183-A2.
XX PN 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX

```



PA (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA27520.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 25; SEQ ID NO 51574; 1766bp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 344 AA;  
 SQ  
 Query Match 2.1%; Score 7; DB 6; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 EIVKILK 22  
 Db 335 EIVKILK 341  
 |||||  
 RESULT 73  
 AAG05089  
 ID AAG05089 standard; protein; 345 AA.  
 XX  
 XX AAG05089;  
 AC  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 1362.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS  
 XX Arabidopsis thaliana.  
 XX

PN EP1033405-A2.  
 XX 06-SEP-2000.  
 PD 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 XX 09-MAR-1999; 99US-0123548P.  
 XX 23-MAR-1999; 99US-0125788P.  
 XX 25-MAR-1999; 99US-0126264P.  
 XX 29-MAR-1999; 99US-0126785P.  
 XX 01-APR-1999; 99US-0127462P.  
 XX 06-APR-1999; 99US-0128234P.  
 XX 08-APR-1999; 99US-0128714P.  
 XX 16-APR-1999; 99US-0129845P.  
 XX 19-APR-1999; 99US-0130077P.  
 XX 21-APR-1999; 99US-0130449P.  
 XX 23-APR-1999; 99US-0130510P.  
 XX 28-APR-1999; 99US-0130891P.  
 XX 30-APR-1999; 99US-0131449P.  
 XX 04-MAY-1999; 99US-0132048P.  
 XX 05-MAY-1999; 99US-0132407P.  
 XX 06-MAY-1999; 99US-0132484P.  
 XX 06-MAY-1999; 99US-0132485P.  
 XX 07-MAY-1999; 99US-0132486P.  
 XX 11-MAY-1999; 99US-0132863P.  
 XX 14-MAY-1999; 99US-0134218P.  
 XX 14-MAY-1999; 99US-0134219P.  
 XX 14-MAY-1999; 99US-0134221P.  
 XX 18-MAY-1999; 99US-0134370P.  
 XX 19-MAY-1999; 99US-0134768P.  
 XX 20-MAY-1999; 99US-0134941P.  
 XX 21-MAY-1999; 99US-0135124P.  
 XX 25-MAY-1999; 99US-0135353P.  
 XX 27-MAY-1999; 99US-0135629P.  
 XX 27-MAY-1999; 99US-0136021P.  
 XX 27-MAY-1999; 99US-0136392P.  
 XX 28-MAY-1999; 99US-0136782P.  
 XX 01-JUN-1999; 99US-0137222P.  
 XX 03-JUN-1999; 99US-0137528P.  
 XX 07-JUN-1999; 99US-0137502P.  
 XX 08-JUN-1999; 99US-0137724P.  
 XX 10-JUN-1999; 99US-0138094P.  
 XX 10-JUN-1999; 99US-0138540P.  
 XX 14-JUN-1999; 99US-0138847P.  
 XX 16-JUN-1999; 99US-0139119P.  
 XX 16-JUN-1999; 99US-0139452P.  
 XX 17-JUN-1999; 99US-0139453P.  
 XX 18-JUN-1999; 99US-0139492P.  
 XX 18-JUN-1999; 99US-0139454P.  
 XX 18-JUN-1999; 99US-0139455P.  
 XX 18-JUN-1999; 99US-0139456P.  
 XX 18-JUN-1999; 99US-0139457P.  
 XX 18-JUN-1999; 99US-0139458P.  
 XX 18-JUN-1999; 99US-0139459P.  
 XX 18-JUN-1999; 99US-0139460P.  
 XX 18-JUN-1999; 99US-0139461P.  
 XX 18-JUN-1999; 99US-0139462P.  
 XX 18-JUN-1999; 99US-0139463P.  
 XX 18-JUN-1999; 99US-0139750P.  
 XX 21-JUN-1999; 99US-0139763P.  
 XX 22-JUN-1999; 99US-0139817P.  
 XX 23-JUN-1999; 99US-0139899P.  
 XX 23-JUN-1999; 99US-0140353P.  
 XX 23-JUN-1999; 99US-0140354P.  
 XX 24-JUN-1999; 99US-0140695P.  
 XX 28-JUN-1999; 99US-0140823P.  
 XX 29-JUN-1999; 99US-0140991P.  
 XX 30-JUN-1999; 99US-0141287P.  
 XX 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142927P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145125P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 345;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 186 TFKLLLT 192  
|||  
Db 190 TFKLLLT 196

RESULT 74  
ABB91447  
ID ABB91447 standard; protein; 368 AA.  
XX  
AC ABB91447;

DT 31-MAY-2002 (first entry)  
DE Herbicidally active polypeptide SEQ ID NO 658.  
KW Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
OS WO200210210-A2.  
PN 07-FEB-2002.  
PD 28-AUG-2001; 2001WO-EP009892.  
PF 28-AUG-2001; 2001WO-EP009892.  
XX (FARB ) BAYER AG.  
PA



PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139811P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144008P.  
PR 16-JUL-1999; 99US-0144088P.  
PR 16-JUL-1999; 99US-0144088P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161614P.

PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 28-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 28-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307  
 |||||  
 Db 112 LIEFLSS 118

RESULT 77  
 ABB92224  
 ID ABB92224 standard; protein; 377 AA.

AC ABB92224;  
 XX  
 DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1435.

DE Herbicidal; plant; agriculture; herbicide.

KW Arabidopsis thaliana.

XX WO200210210-A2.

PN 07-FEB-2002.

PD 28-AUG-2001; 2001WO-EP009892.

PF 28-AUG-2001; 2001WO-EP009892.

PR (FARB ) BAYER AG.

PA Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.

PS Claim 5; SEQ ID NO 1435; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides

SQ Sequence 377 AA;

Query Match 2.1%; Score 7; DB 5; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307  
 |||||  
 Db 112 LIEFLSS 118

RESULT 78  
 AAB85255

ID AAB85255 standard; protein; 384 AA.

XX AAB85255;

XX 07-SEP-2001 (first entry)

XX Plant porphobilinogen synthase (PPS).

XX Thioedoxin; porphobilinogen synthase; T-PPS; PPS; enzyme; herbicide;  
 KM delta-aminolevulinic acid; plant growth inhibitor.

XX Lycopersicon esculentum.

XX WO200146446-A1.

XX 28-JUN-2001.

XX 19-DEC-2000; 2000WO-US034584.

XX 22-DEC-1999; 99US-0171785P.

XX (PARA-) PARADIGM GENETICS INC.

XX Crawford JM, Rice J, Sevala V, Stewart S;

XX WPI; 2001-418081/44.

XX N-PSDB; AAH22805.

XX Novel plant thioedoxin-porphobilinogen synthase or porphobilinogen  
 PT synthase polypeptides, useful for identifying compounds for use as  
 PT herbicides by inhibiting enzymatic activity of the polypeptides.

XX Claim 4; Page 21-22; 25pp; English.

XX The invention provides novel DNA sequences encoding enzymes such as plant  
 CC thioedoxin-porphobilinogen synthase (T-PPS) and plant porphobilinogen  
 CC synthase (PPS). Methods of determining the enzymatic activity of T-PPS or  
 CC PPS or its functional fragment are provided that involves contacting  
 CC delta-aminolevulinic acid with the protein or its functional fragment and  
 CC measuring the amount of porphobilinogen formed from it. Compounds which  
 CC can modify the enzymatic activity T-PPS or PPS can also be identified  
 CC similarly, which are useful for inhibiting plant growth by inhibiting  
 CC enzymatic activity of T-PPS or PPS or its functional fragment. The  
 CC compounds thus identified are useful as herbicides. The present sequence  
 CC represents the PPS enzyme

XX Sequence 384 AA;

Query Match 2.1%; Score 7; DB 4; Length 384;  
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LLRDKSP 267  
 |||||  
 Db 316 LLRDKSP 322

RESULT 79  
 AAG51052

ID AAG51052 standard; protein; 400 AA.

XX AAG51052;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 64757.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;

termination sequence.  
Arabidopsis thaliana.  
EP1033405-A2.  
06-SEP-2000.  
25-FEB-2000; 2000EP-00301439.  
25-FEB-1999; 99US-0121825P.  
05-MAR-1999; 99US-0123180P.  
09-MAR-1999; 99US-0123548P.  
23-MAR-1999; 99US-0125788P.  
25-MAR-1999; 99US-0126264P.  
29-MAR-1999; 99US-0126785P.  
01-APR-1999; 99US-0127462P.  
06-APR-1999; 99US-0128234P.  
08-APR-1999; 99US-0128714P.  
16-APR-1999; 99US-0129845P.  
19-APR-1999; 99US-0130077P.  
21-APR-1999; 99US-0130449P.  
23-APR-1999; 99US-0130510P.  
23-APR-1999; 99US-0130891P.  
28-APR-1999; 99US-0131449P.  
30-APR-1999; 99US-0132407P.  
04-MAY-1999; 99US-0132484P.  
05-MAY-1999; 99US-0132485P.  
06-MAY-1999; 99US-0132486P.  
07-MAY-1999; 99US-0132487P.  
07-MAY-1999; 99US-0132487P.  
11-MAY-1999; 99US-0134256P.  
14-MAY-1999; 99US-0134218P.  
14-MAY-1999; 99US-0134219P.  
14-MAY-1999; 99US-0134221P.  
14-MAY-1999; 99US-0134370P.  
18-MAY-1999; 99US-0134768P.  
19-MAY-1999; 99US-0134941P.  
20-MAY-1999; 99US-0135124P.  
21-MAY-1999; 99US-0135353P.  
24-MAY-1999; 99US-0135629P.  
25-MAY-1999; 99US-0136021P.  
27-MAY-1999; 99US-0136392P.  
28-MAY-1999; 99US-0136782P.  
01-JUN-1999; 99US-0137222P.  
03-JUN-1999; 99US-0137528P.  
04-JUN-1999; 99US-0137502P.  
07-JUN-1999; 99US-0137724P.  
08-JUN-1999; 99US-0138094P.  
10-JUN-1999; 99US-0138540P.  
10-JUN-1999; 99US-0138847P.  
14-JUN-1999; 99US-0139119P.  
16-JUN-1999; 99US-0139452P.  
17-JUN-1999; 99US-0139453P.  
18-JUN-1999; 99US-0139492P.  
18-JUN-1999; 99US-0139454P.  
18-JUN-1999; 99US-0139455P.  
18-JUN-1999; 99US-0139456P.  
18-JUN-1999; 99US-0139457P.  
18-JUN-1999; 99US-0139458P.  
18-JUN-1999; 99US-0139459P.  
18-JUN-1999; 99US-0139460P.  
18-JUN-1999; 99US-0139461P.  
18-JUN-1999; 99US-0139463P.  
18-JUN-1999; 99US-0139750P.  
18-JUN-1999; 99US-0139763P.  
21-JUN-1999; 99US-0139817P.  
22-JUN-1999; 99US-0139899P.  
23-JUN-1999; 99US-0140353P.  
23-JUN-1999; 99US-0140354P.  
24-JUN-1999; 99US-0140695P.  
28-JUN-1999; 99US-0140823P.  
29-JUN-1999; 99US-0140891P.  
30-JUN-1999; 99US-0141287P.  
01-JUL-1999; 99US-0141842P.  
01-JUL-1999; 99US-0142154P.  
02-JUL-1999; 99US-0142055P.  
06-JUL-1999; 99US-0142390P.  
08-JUL-1999; 99US-0142803P.  
09-JUL-1999; 99US-0142920P.  
12-JUL-1999; 99US-0142977P.  
13-JUL-1999; 99US-0143542P.  
14-JUL-1999; 99US-0143624P.  
15-JUL-1999; 99US-0144005P.  
16-JUL-1999; 99US-0144085P.  
19-JUL-1999; 99US-0144086P.  
19-JUL-1999; 99US-0144325P.  
19-JUL-1999; 99US-0144331P.  
19-JUL-1999; 99US-0144332P.  
19-JUL-1999; 99US-0144333P.  
19-JUL-1999; 99US-0144334P.  
19-JUL-1999; 99US-0144335P.  
20-JUL-1999; 99US-0144352P.  
20-JUL-1999; 99US-0144632P.  
20-JUL-1999; 99US-0144884P.  
21-JUL-1999; 99US-0144814P.  
21-JUL-1999; 99US-0145086P.  
21-JUL-1999; 99US-0145088P.  
22-JUL-1999; 99US-0145085P.  
22-JUL-1999; 99US-0145087P.  
22-JUL-1999; 99US-0145089P.  
23-JUL-1999; 99US-0145192P.  
23-JUL-1999; 99US-0145145P.  
23-JUL-1999; 99US-0145218P.  
23-JUL-1999; 99US-0145224P.  
26-JUL-1999; 99US-0145276P.  
27-JUL-1999; 99US-0145913P.  
27-JUL-1999; 99US-0145918P.  
28-JUL-1999; 99US-0145919P.  
28-JUL-1999; 99US-0145951P.  
02-AUG-1999; 99US-0146386P.  
02-AUG-1999; 99US-0146388P.  
02-AUG-1999; 99US-0146389P.  
03-AUG-1999; 99US-0147038P.  
04-AUG-1999; 99US-0147204P.  
04-AUG-1999; 99US-0147302P.  
05-AUG-1999; 99US-0147192P.  
05-AUG-1999; 99US-0147260P.  
06-AUG-1999; 99US-0147303P.  
06-AUG-1999; 99US-0147416P.  
09-AUG-1999; 99US-0147493P.  
09-AUG-1999; 99US-0147935P.  
10-AUG-1999; 99US-0148171P.  
11-AUG-1999; 99US-0148319P.  
12-AUG-1999; 99US-0148341P.  
13-AUG-1999; 99US-0148565P.  
16-AUG-1999; 99US-0148684P.  
16-AUG-1999; 99US-0149368P.  
17-AUG-1999; 99US-0149175P.  
18-AUG-1999; 99US-0149426P.  
20-AUG-1999; 99US-0149722P.  
20-AUG-1999; 99US-0149723P.  
20-AUG-1999; 99US-0149929P.  
23-AUG-1999; 99US-0149902P.  
23-AUG-1999; 99US-0149930P.  
25-AUG-1999; 99US-0150566P.  
26-AUG-1999; 99US-0150884P.  
27-AUG-1999; 99US-0151065P.  
27-AUG-1999; 99US-0151066P.  
27-AUG-1999; 99US-0151080P.  
30-AUG-1999; 99US-0151303P.  
31-AUG-1999; 99US-0151438P.  
01-SEP-1999; 99US-0151930P.  
07-SEP-1999; 99US-0152363P.

```

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158923P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      2.1%; Score 7; DB 3; Length 400;
Best Local Similarity 100.0%; Pred. No. 6e+02; 0; Indels
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 272 EAFHVPK 278
Db 335 EAFHVPK 341

RESULT 80
AAG39773
ID AAG39773 standard; protein; 422 AA.
XX
AC AAG39773;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49262.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.

```



```
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142922P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145222P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0145388P.
PR 02-AUG-1999; 99US-0145388P.
PR 02-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148315P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149168P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149428P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149932P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.

PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158039P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 422;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGLL 79
Db 69 LYSSGLL 75

RESULT 81
AAE24493
ID AAE24493 standard; protein; 422 AA.
XX
AC AAE24493;
XX
DT 04-OCT-2002 (first entry)
XX
DE Human RATL1d6 (regulated in activated T-lymphocyte 1d6) protein.
XX
KW Human; ubiquitin conjugating enzyme; UBC; RATL1d6; immune disorder;
KW regulated in activated T-lymphocyte 1d6; neuronal disorder; cancer;
KW tumour; lymphoproliferative; cancer; adenocarcinoma; leukaemia; myeloma;
KW sarcoma; neurodegenerative; inflammatory; rheumatoid arthritis; asthma;
KW multiple sclerosis; psoriasis; neuronal; Alzheimer's disease; dementia;
KW depression; epilepsy; acquired immuno deficiency syndrome; allergy; AIDS;
KW anaemia; atopic dermatitis; diabetes mellitus; dermatological;
KW myocardial infarction; renal tubular acidosis; gonadal dysgenesis;
KW dysplasia; cataract; cytostatic; neuroprotective; nootropic; anti-HIV;
KW anticonvulsant; antiinflammatory; Cushing's syndrome; cardiact;
KW ophthalmological.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 89. .333
```



XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA35658.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 59712; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 439 AA;  
  
Query Match 2.1%; Score 7; DB 6; Length 439;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 240 HNFAMT 246  
DB 117 HNFAMT 123  
  
RESULT 84  
ABB52729  
ID ABB52729 standard; protein; 452 AA.  
XX  
XX ABB52729;  
AC  
XX  
XX 11-FEB-2002 (first entry)

XX Escherichia coli polypeptide SEQ ID NO 859.  
XX  
XX  
KW Escherichia coli; B2/D+A; antiinflammatory; antibacterial;  
KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;  
KW systemic infection; non-diarrhoeal infection; septicemia;  
KW pyelonephritis; antibiotic resistance.  
XX  
OS Escherichia coli.  
XX  
XX WO200166572-A2.  
PD  
PD 13-SEP-2001.  
XX  
XX 12-MAR-2001; 2001WO-EP003445.  
XX  
XX 10-MAR-2000; 2000FR-00003145.  
PR 02-FEB-2001; 2001FR-00001449.  
XX  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;  
PI WPI; 2001-550253/61.  
XX  
XX A library of DNA fragments of Escherichia coli strains for the phylogenetic  
PT determination of a given strain comprises polynucleotides of nature B2/D+  
PT A-.  
XX  
XX Example 6; Fig 6; 646pp; English.  
XX  
XX The invention relates to a library of DNA fragments of Escherichia coli  
CC strains comprising polynucleotides (ABA985177-ABA98729 and ABA89533) and  
CC encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature  
CC B2/D+A-. The polynucleotides have potential antiinflammatory,  
CC antibacterial and immunosuppressive activity as part of pharmaceutical  
CC compositions used to treat, palliate or prevent extra-intestinal *E. coli*  
CC infections. The polypeptides are useful for determining the phylogenetic  
CC group of a given *E. coli* strain. These polypeptides can detect and treat  
CC an undesired development of *E. coli*, particularly an extra-intestinal  
CC infection that include systemic and non-diarrhoeal infections such as  
CC septicemia, pyelonephritis and meningitis this is particularly  
CC advantageous as bacterial resistance is increasing with the more frequent  
CC use of broad spectrum antibiotics  
XX  
SQ Sequence 452 AA;  
  
Query Match 2.1%; Score 7; DB 4; Length 452;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 178 DIASDAF 184  
DB 128 DIASDAF 134  
  
RESULT 85  
ABP73473  
ID ABP73473 standard; protein; 458 AA.  
XX  
XX ABP73473;  
AC  
XX 30-JAN-2003 (first entry)  
XX  
XX Candida albicans essential protein SEQ ID NO 7310.  
XX  
XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
KW signal transduction; DNA replication; cell division; growth;  
KW proliferation; Candida albicans; fungicide; antifungal.  
XX  
XX Candida albicans.  
XX  
XX WO200253728-A2.  
XX  
XX

XX 11-JUL-2002.  
 XX PD  
 XX PF 26-DEC-2001; 2001WO-US049486.  
 XX PR 29-DEC-2000; 2000US-0259128P.  
 XX PR 20-FEB-2001; 2001US-00792024.  
 XX PR 22-AUG-2001; 2001US-0314050P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
 XX DR WPI: 2002-566694/60.  
 XX DR N-PSDB; ABZ32023.  
 XX PT Constructing strains for identifying gene products as effective targets  
 XX PT for therapeutic intervention, by inactivating in the strain one allele of  
 XX PT a gene and placing other allele of the gene under conditional expression.  
 XX PS Claim 44; SEQ ID NO 7310; 167pp + Sequence Listing; English.  
 XX CC The invention relates to constructing (M1) a strain of diploid fungal  
 CC cells in which both alleles of a gene are modified, comprising modifying  
 CC one allele by insertion or replacement by a cassette having an  
 CC expressible selectable marker and modifying other allele by  
 CC recombination, of a promoter replacement fragment with a heterologous  
 CC promoter, so that expression of the second allele is regulated by the  
 CC promoter. (M1) is useful for constructing a strain of diploid fungal  
 CC cells in which both alleles of a gene are modified. The diploid fungal  
 CC cells having both alleles modified are useful for identifying a gene that  
 CC is essential to the survival or growth of a fungus, a gene that  
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene  
 CC that contributes to the resistance of a diploid fungus to an antifungal  
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
 CC and for identifying a therapeutic agent for treatment of a mammalian  
 CC disease. (M1) is useful for identifying a compound which modulates the  
 CC activity of a gene product, preferably enzymatic activity, carbon  
 CC compound catabolism, biosynthetic, transporter, transcriptional,  
 CC translational, signal transduction, DNA replication and cell division  
 CC activity. The method is useful for identifying a compound having the  
 CC ability to inhibit growth or proliferation of C. albicans cells and for  
 CC treating infection by C. albicans. The present sequence is that of an  
 CC essential Candida albicans protein used in the method of the invention.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification but is based on sequence information supplied to Derwent by  
 CC the European Patent Office  
 XX SQ Sequence 458 AA;  
 Query Match 2.1%; Score 7; DB 5; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 318 FADEKXY 324  
 Db 332 FADEKXY 338  
 RESULT 86  
 AAG32191  
 ID AAG32191 standard; protein; 482 AA.  
 XX AC AAG32191;  
 XX DT 17-OCT-2000 (first entry)  
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38786.  
 XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 OS Arabidopsis thaliana.  
 XX EP1033405-A2.  
 XX 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 XX 09-MAR-1999; 99US-0123548P.  
 XX 23-MAR-1999; 99US-0125788P.  
 XX 25-MAR-1999; 99US-0126264P.  
 XX 29-MAR-1999; 99US-0126785P.  
 XX 01-APR-1999; 99US-0127462P.  
 XX 06-APR-1999; 99US-0128234P.  
 XX 08-APR-1999; 99US-0128714P.  
 XX 16-APR-1999; 99US-0129845P.  
 XX 19-APR-1999; 99US-0130077P.  
 XX 21-APR-1999; 99US-0130449P.  
 XX 23-APR-1999; 99US-0130510P.  
 XX 28-APR-1999; 99US-0130891P.  
 XX 30-APR-1999; 99US-0131449P.  
 XX 30-APR-1999; 99US-0132048P.  
 XX 04-MAY-1999; 99US-0132407P.  
 XX 05-MAY-1999; 99US-0132484P.  
 XX 06-MAY-1999; 99US-0132485P.  
 XX 06-MAY-1999; 99US-0132486P.  
 XX 07-MAY-1999; 99US-0132487P.  
 XX 07-MAY-1999; 99US-0132863P.  
 XX 11-MAY-1999; 99US-0134256P.  
 XX 14-MAY-1999; 99US-0134218P.  
 XX 14-MAY-1999; 99US-0134219P.  
 XX 14-MAY-1999; 99US-0134221P.  
 XX 14-MAY-1999; 99US-0134370P.  
 XX 18-MAY-1999; 99US-0134768P.  
 XX 19-MAY-1999; 99US-0134941P.  
 XX 20-MAY-1999; 99US-0135124P.  
 XX 21-MAY-1999; 99US-0135353P.  
 XX 24-MAY-1999; 99US-0135639P.  
 XX 25-MAY-1999; 99US-0136021P.  
 XX 27-MAY-1999; 99US-0136392P.  
 XX 28-MAY-1999; 99US-0136782P.  
 XX 01-JUN-1999; 99US-0137222P.  
 XX 03-JUN-1999; 99US-0137528P.  
 XX 04-JUN-1999; 99US-0137502P.  
 XX 07-JUN-1999; 99US-0137724P.  
 XX 08-JUN-1999; 99US-0138094P.  
 XX 10-JUN-1999; 99US-0138540P.  
 XX 14-JUN-1999; 99US-0139119P.  
 XX 16-JUN-1999; 99US-0139452P.  
 XX 16-JUN-1999; 99US-0139453P.  
 XX 17-JUN-1999; 99US-0139492P.  
 XX 18-JUN-1999; 99US-0139454P.  
 XX 18-JUN-1999; 99US-0139455P.  
 XX 18-JUN-1999; 99US-0139456P.  
 XX 18-JUN-1999; 99US-0139457P.  
 XX 18-JUN-1999; 99US-0139458P.  
 XX 18-JUN-1999; 99US-0139459P.  
 XX 18-JUN-1999; 99US-0139460P.  
 XX 18-JUN-1999; 99US-0139461P.  
 XX 18-JUN-1999; 99US-0139462P.  
 XX 18-JUN-1999; 99US-0139463P.  
 XX 18-JUN-1999; 99US-0139750P.  
 XX 18-JUN-1999; 99US-0139763P.  
 XX 21-JUN-1999; 99US-0139817P.  
 XX 22-JUN-1999; 99US-0139899P.  
 XX 23-JUN-1999; 99US-0140353P.  
 XX 23-JUN-1999; 99US-0140354P.  
 XX 24-JUN-1999; 99US-0140695P.  
 XX 28-JUN-1999; 99US-0140823P.  
 XX 29-JUN-1999; 99US-0140991P.

```

PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 482;
Best Local Similarity 100.0%; Pred. No. 7.le+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIEFLSS 307
DB 100 LIEFLSS 106

RESULT 87
AAV49289
ID AAV49289 standard; protein; 490 AA.
XX
AC AAV49289;
XX
DT 07-FEB-2000 (first entry)
XX
DE Mouse GLCIA polypeptide.
XX
KW Glaucoma; PCR amplification; primary open wide angle glaucoma;
KW GLCIA gene; mouse.
XX
OS Mus sp.
XX
PN WO9951779-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007671.
XX
```



```

CC  ichthyosis
XX
SQ  Sequence 492 AA;

Query Match      2.1%; Score 7; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  45 SLOAMKE 51
DB  282 SLOAMKE 288
|||||

RESULT 90
AAB85155
ID  AAB85155 standard; protein; 492 AA.
XX
AC  AAB85155;
XX
DT  11-SEP-2003 (revised)
DT  22-AUG-2001 (first entry)
XX
DE  Zebrafish cytochrome P450 polypeptide, P450RAI.
XX
KW  Cytochrome P450; P450RAI-2; brain; retinoic acid; cancer; dysplasia;
KW  autoimmune; dermatological; cytotatic; antiinflammatory; antiseborrheic;
KW  antipsoriatic; immunosuppressive.
XX
OS  Danio rerio.
XX
PN  WO200144443-A2.
XX
PD  21-JUN-2001.
XX
PF  15-DEC-2000; 2000WO-CA001493.
XX
PR  16-DEC-1999; 99US-0171110P.
PR  27-JAN-2000; 2000US-0178314P.
XX
PA  (CYTO-) CYTOCHROMA INC.
XX
PI  White JA, Petkovich PM, Jones G, Ramshaw H;
XX
DR  WPI; 2001-390242/41.
DR  N-PSDB; AAB22442.
XX
PT  Novel P450 protein useful for metabolizing retinoic acid for treating
PT  cancer, dysplasia, an autoimmune or dermatological disease.
XX
PS  Claim 106; Page 164-166; 174pp; English.
XX
CC  The present invention provides a novel all-trans-RA metabolising
CC  cytochrome P450, P450RAI-2, that is predominantly expressed in the brain.
CC  This novel cytochrome P450 is useful for metabolising retinoic acid in a
CC  cell or organism, for screening drugs for their effect of protein
CC  activity, oxidizing a retinoid, screening an agent for its effect on
CC  protein activity. The P450RAI-2 polypeptide, nucleic acid and host cells
CC  containing them are useful for treating cancer, dysplasia, an autoimmune
CC  or dermatological disease. A drug which has an effect on the expression
CC  of P450RAI-2 is used to inhibit retinoic acid metabolism in the treatment
CC  cancer, actinic keratosis, oral leukoplakia, a secondary head and/or neck
CC  tumour, a non-small cell lung carcinoma, a basal cell carcinoma, skin
CC  cancer, and a premalignancy associated actinic keratosis, acne,
CC  psoriasis, ichthyosis, and/or preferably acute promyelocytic leukemia.
CC  The present sequence represents a zebrafish P450RAI protein that
CC  hydroxylates retinoic acid at the 4 position of the beta-ionone ring.
CC  (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ  Sequence 492 AA;

Query Match      2.1%; Score 7; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  45 SLOAMKE 51
DB  282 SLOAMKE 288
|||||

RESULT 91
ABP52146
ID  ABP52146 standard; protein; 492 AA.
XX
AC  ABP52146;
XX
DT  29-AUG-2003 (revised)
DT  13-OCT-2002 (first entry)
XX
DE  Zebrafish P450RAI-2 homologous protein sequence SEQ ID NO:14.
XX
KW  Cytochrome P450; dermatological disorder; cancer; brain disorder;
KW  cytotatic; immunosuppressive; dermatological; antisense therapy;
KW  P450RAI-2; inhibiting P450RAI-2 induced retinoic acid hydroxylation;
KW  actinic keratosis; oral leukoplakia; tumour; basal cell carcinoma;
KW  non-small cell lung carcinoma; acute promyelocytic leukaemia; acne;
KW  psoriasis; ichthyosis.
XX
OS  Danio rerio.
XX
PN  WO200248334-A2.
XX
PD  20-JUN-2002.
XX
PF  17-DEC-2001; 2001WO-CA001805.
XX
PR  15-DEC-2000; 2000WO-CA001493.
XX
PA  (CYTO-) CYTOCHROMA INC.
XX
PI  White JA, Petkovich PM, Jones G, Ramshaw H;
XX
DR  WPI; 2002-583506/62.
DR  N-PSDB; ABQ74193.
XX
PT  Novel polyclonal antibody specific to human cytochrome P450 retinoic acid
PT  metabolizing protein, P450RAI-2, useful for inhibiting P450RAI-2 induced
PT  retinoic acid hydroxylation in a human being treated for cancer.
XX
PS  Example 1; Page 164-166; 179pp; English.
XX
CC  The present invention describes a polyclonal antibody (I) to a human
CC  cytochrome P450 retinoic acid metabolising peptide (P450RAI-2) comprising
CC  a sequence (see ABP52142) of 512 amino acids (I) has cytotatic,
CC  immunosuppressive and dermatological activities, and can be used in
CC  antisense therapy. (I) can be used for inhibiting P450RAI-2 induced
CC  retinoic acid hydroxylation in an organism, in particular a human being
CC  treated for a disease such as cancer, actinic keratosis, oral
CC  leukoplakia, secondary tumour of the head and/or neck, non-small cell
CC  lung carcinoma, basal cell carcinoma, acute promyelocytic leukaemia,
CC  lung, skin cancer and pre-malignancy associated actinic keratosis, acne,
CC  psoriasis and/or ichthyosis, or an in vitro system. (I) is useful for
CC  screening for the expression of P450RAI-2 in a sample, where the antibody
CC  is labeled to enable detection of binding and non-binding to a P450RAI-2
CC  substrate and the antibody interaction is detected by an ELISA assay.
CC  This method is useful for diagnosing non small lung cell carcinoma in a
CC  patient. The present sequence represents a zebrafish protein which is
CC  homologous to the human P450RAI-2 sequence, which is given in an example
CC  from the present invention. (Updated on 29-AUG-2003 to standardise OS
CC  field)
XX
SQ  Sequence 492 AA;

Query Match      2.1%; Score 7; DB 5; Length 492;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



QY 45 SLOANKE 51  
 Db 282 SLOANKE 288  
 RESULT 92  
 AAE15325  
 ID AAE15325 standard; protein; 492 AA.  
 XX AC AAE15325;  
 XX AC  
 DT 29-AUG-2003 (revised)  
 DT 07-MAR-2002 (first entry)  
 XX Zebrafish P450RAI protein.  
 DE Zebrafish; retinoid metabolism; retinoic acid; RA; haeme-binding motif;  
 KW vitamin A; cytochrome P450; prostate cancer; drug screening; P450RAI.  
 KW OS Danio rerio.  
 XX US6306624-B1.  
 XX 23-OCT-2001.  
 XX 25-JUN-1997; 97US-00882164.  
 XX 21-JUN-1996; 96US-00667546.  
 PR 01-OCT-1996; 96US-00724466.  
 PR 23-JUN-1997; 97WO-CA000440.  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 PA Petkovich PM, White JA, Beckett BR, Jones G;  
 PI WPI; 2002-033254/04.  
 DR N-PSDB; AAD24484.  
 XX New DNA fragments having promoter activity, useful in retinoid  
 PT metabolism, as well as in producing retinoic acid metabolizing cytochrome  
 PT P450s that are useful as targets for the treatment of certain cancers.  
 XX Example 1; Fig 2C; 75pp; English.  
 XX The present invention relates to retinoid (e.g., retinoic acid (RA),  
 CC vitamin A) metabolising proteins and nucleic acid sequences encoding  
 CC them. RA metabolising proteins contain a haeme-binding motif which is  
 CC characteristic of the group of proteins known as cytochrome P450s. The  
 CC sequences of the invention are useful in retinoid metabolism and in  
 CC producing retinoic acid metabolising cytochrome P450s. They are  
 CC particularly useful as targets for the treatment of certain cancers such  
 CC as prostate cancer. The invention also relates to a method of screening  
 CC drugs for their effect on activity of RA inducible proteins. The present  
 CC sequence is P450RAI protein from zebrafish. (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 XX Sequence 492 AA;  
 SQ Query Match 2.1%; Score 7; DB 5; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 45 SLOANKE 51  
 Db 282 SLOANKE 288  
 RESULT 93  
 AAR98225  
 ID AAR98225 standard; protein; 497 AA.  
 XX AC AAR98225;  
 XX

DT 23-SEP-1996 (first entry)  
 XX Trabecular meshwork induced glucocorticoid response protein.  
 DE Trabecular meshwork induced glucocorticoid response protein; TIGR;  
 XX glaucoma; therapy; diagnosis.  
 KW Homo sapiens.  
 XX OS  
 XX FH Key Location/Qualifiers  
 FT Modified-site 57..60  
 FT /label= N-glycosylation\_site  
 FT Region 85..92  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Region 92..99  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Binding-site 110..113  
 FT /label= Heparin\_sulphate\_binding\_site  
 FT Region 121..128  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Region 128..135  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Region 135..142  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Region 142..149  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Region 142..149  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Binding-site 146..150  
 FT /label= Heparin\_sulphate\_binding\_site  
 FT Region 149..159  
 FT /label= Consensus\_leucine\_zipper\_unit  
 FT Modified-site 221..222  
 FT /label= O-glycosylation\_site  
 FT Modified-site 222..223  
 FT /label= O-glycosylation\_site  
 FT Domain 223..224  
 FT /label= Heparin\_sulphate\_initiation\_domain  
 FT Domain 231..232  
 FT /label= Heparin\_sulphate\_initiation\_domain  
 FT Modified-site 270..272  
 FT /label= O-glycosylation\_site  
 FT Modified-site 305..306  
 FT /label= O-glycosylation\_site  
 FT Domain 324..325  
 FT /label= Heparin\_sulphate\_initiation\_domain  
 FT Modified-site 397..401  
 FT /label= O-glycosylation\_site  
 FT Modified-site 453..457  
 FT /label= O-glycosylation\_site  
 FT Domain 457..459  
 FT /label= O-glycosylation\_site  
 XX WO9614411-A1.  
 XX 17-MAY-1996.  
 XX 27-OCT-1995; 95WO-US014024.  
 XX 03-NOV-1994; 94US-00336235.  
 XX (REGC ) UNIV CALIFORNIA.  
 XX Nguyen TD, Polansky JR, Huang W;  
 DR WPI; 1996-251761/25.  
 DR N-PSDB; AAT30152.  
 XX Trabecular meshwork induced glucocorticoid response protein - useful to  
 PT treat glaucoma and diagnose glaucoma, steroid sensitivity and related  
 FT diseases.  
 XX Claim 1; Fig 1A-1D; 59pp; English.  
 PS

XX A novel trabecular meshwork induced glucocorticoid response protein  
 CC (AAR98225), TIGR, is highly induced by glucocorticoids in the endothelial  
 CC lining of the human trabecular meshwork. Excessive levels of TIGR are  
 CC indicative of glaucoma. TIGR can be obtained by expressing TIGR-encoding  
 CC cDNA (see also AAT30152-53) in a bacterial or eukaryotic host. It can be  
 CC used to raise antibodies useful in the diagnosis of glaucoma, steroid  
 CC sensitivity and related diseases. Leucine zipper-contg. peptides of TIGR  
 CC are also useful diagnostics  
 XX

SQ Sequence 497 AA;

Query Match 2.1%; Score 7; DB 2; Length 497;  
 Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 134 NLLRDKS 140  
 |||||

RESULT 94  
 AAW61390  
 ID AAW61390 standard; protein; 497 AA.  
 XX  
 AC AAW61390;

XX 02-OCT-1998 (first entry)

XX Glucocorticoid-induced protein TIGR.

XX Glucocorticoid-induced protein; TIGR; trabecular meshwork; glaucoma.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Modified-site 57..60  
 FT Region /note= "N-glycosylation site"  
 FT Region 85..92  
 FT Region /note= "Leucine zipper unit"  
 FT Region 92..99  
 FT Domain /note= "Leucine zipper unit"  
 FT Domain 110..113  
 FT Region /note= "Heparin sulfate binding domain"  
 FT Region 121..128  
 FT Region /note= "Leucine zipper unit"  
 FT Region 128..135  
 FT Region /note= "Leucine zipper unit"  
 FT Region 135..142  
 FT Region /note= "Leucine zipper unit"  
 FT Region 142..149  
 FT Domain /note= "Leucine zipper unit"  
 FT Domain 146..150  
 FT Region /note= "Heparin sulfate binding domain"  
 FT Region 149..156  
 FT Modified-site /note= "Leucine zipper unit"  
 FT Modified-site 221..222  
 FT Modified-site /note= "O-glycosylation site"  
 FT Modified-site 222..223  
 FT Domain /note= "O-glycosylation site"  
 FT Domain 223..224  
 FT Domain /note= "Heparin sulfate initiation domain"  
 FT Domain 231..232  
 FT Modified-site /note= "Heparin sulfate initiation domain"  
 FT Modified-site 270..272  
 FT Modified-site /note= "O-glycosylation site"  
 FT Modified-site 305..306  
 FT Domain /note= "O-glycosylation site"  
 FT Domain 324..325  
 FT Modified-site /note= "Heparin sulfate initiation domain"  
 FT Modified-site 397..401  
 FT Modified-site /note= "O-glycosylation site"  
 FT Modified-site 453..457

FT Modified-site /note= "O-glycosylation site"  
 FT 457..459  
 FT /note= "O-glycosylation site"  
 XX  
 PN US5789169-A.  
 XX  
 PD 04-AUG-1998.  
 XX  
 XX 17-MAY-1996; 96US-00649432.  
 XX  
 PR 03-NOV-1994; 94US-00336235.  
 PR 20-OCT-1995; 95US-00546568.  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 PI Polansky JR, Huang W, Nguyen TD;  
 XX  
 DR WPI; 1998-446069/38.  
 DR N-PSDB; AAV28331.  
 XX  
 XX Diagnosis of glaucoma or steroid sensitivity - by protein detection or  
 PT polymorphism analysis.  
 PT  
 PS Claim 67; Fig 1; 26pp; English.  
 XX  
 CC The glucocorticoid-induced protein, TIGR is produced by cells of the  
 CC trabecular meshwork. By determining if the amount of TIGR present in the  
 CC trabecular meshwork of an eye of a patient exceeds the amount of the  
 CC protein present in the trabecular meshwork of an eye of an individual who  
 CC does not have, and is not predisposed to have glaucoma, you can diagnose  
 CC the presence of glaucoma  
 XX  
 SQ Sequence 497 AA;

Query Match 2.1%; Score 7; DB 2; Length 497;  
 Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 134 NLLRDKS 140  
 |||||

RESULT 95  
 AAG39772  
 ID AAG39772 standard; protein; 497 AA.  
 XX  
 AC AAG39772;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 49261.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 23-MAR-1999; 99US-0123548P.

XX 25-MAR-1999; 99US-0125788P.

XX 29-MAR-1999; 99US-0126264P.

XX 01-APR-1999; 99US-0126785P.

XX 06-APR-1999; 99US-0127462P.



PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159638P.  
 PR 21-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160786P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 22-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 25-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 497;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGLL 79  
 DB 144 LYSSGLL 150  
 |||||

RESULT 96  
 ABB65679  
 ID ABB65679 standard; protein; 500 AA.  
 AC ABB65679;  
 XX  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 23829.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL09782.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 23829; 21pp + Sequence Listing; English.  
 XX

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL1840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 500 AA;

Query Match 2.1%; Score 7; DB 4; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84  
 DB 124 LLVTLIA 130  
 |||||

RESULT 97  
 ABU07341  
 ID ABU07341 standard; protein; 500 AA.  
 XX  
 AC ABU07341;  
 XX  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE Human TIGR/Myocilin protein.  
 XX  
 KW Human; TIGR; MYOC; Myocilin; Glaucoma; blindness;  
 KW trabecular meshwork inducible glucocorticoid responsive protein;  
 KW retinal degenerative disease; RDP; retinitis pigmentosa;  
 KW macular degeneration; Usher syndrome; cardiovascular disease;  
 KW congenital heart disease; myocardial ischaemia; stroke;  
 KW acute endocarditis; hypertensive heart disease; arrhythmia;  
 KW arteriosclerotic heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200282969-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 11-DEC-2001; 2001WO-US048622.  
 XX  
 PR 05-APR-2001; 2001US-0281442P.  
 PR 23-JUL-2001; 2001US-0306889P.  
 XX  
 PA (KONG/) KONG T H.  
 XX  
 PI Kong TH;  
 XX  
 DR WPI; 2003-058597/05.  
 DR N-PSDB; ABX10137, ABX10138.  
 XX  
 PT Determining the presence or the risk of having glaucoma, retinal  
 PT degenerative or cardiovascular diseases in a subject, comprises  
 PT generating transcriptional or translational profiles based on myocilin  
 PT nucleic acids and proteins.  
 XX  
 PS Claim 18; Fig 3; 55pp; English.  
 XX

CC The invention relates to determining whether a subject has or is at risk  
 CC of developing glaucoma, retinal degenerative disease, or a cardiovascular  
 CC disease, comprises generating a transcriptional or translational profile  
 CC (i.e. 'fingerprint') in the subject or in a sample obtained from the  
 CC subject, based on the expression of the different myocilin (MYOC, also  
 CC known as trabecular meshwork inducible glucocorticoid responsive protein,  
 CC TIGR) mRNA species or polypeptide forms, where a difference in the  
 CC profile relative to that in a normal subject indicates that the subject

The invention discloses a composition comprising two or more isolated rat

Claim 48; Fig 8; 105pp; English.

XX This sequence represents a novel human trabecular meshwork induced  
 CC glucocorticoid response protein (TIGR) which is used in a method for  
 CC diagnosing glaucoma in a patient. The method involves the detection of  
 CC polymorphisms whose presence is predictive of a mutation affecting TIGR  
 CC response in the patient and can be diagnostic of glaucoma or steroid  
 CC sensitivity. Base substitutions and base additions upstream of and within  
 CC TIGR exons can also be used to diagnose glaucoma

XX SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 100  
 AAW60670  
 ID AAW60670 standard; protein; 504 AA.  
 XX AC AAW60670;  
 XX DT 14-SEP-1998 (first entry)  
 XX DE Human glaucoma associated protein GLCIA.  
 XX KW Glaucoma; GLCIA; treatment; mutant; juvenile open angle glaucoma; JOAG.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 178 /label= unknown  
 FT /note= "encoded by NCA"

XX PN WO9820131-A1.  
 XX 14-MAY-1998.  
 XX 07-NOV-1997; 97WO-US020702.  
 XX 08-NOV-1996; 96US-00748479.  
 XX 30-JAN-1997; 97US-00791347.  
 XX 21-MAR-1997; 97US-00822999.  
 XX (IOWA ) UNIV IOWA RES FOUND.  
 XX Stone EM, Sheffield V, Alward WLM;  
 XX MPI; 1998-286947/25.  
 XX N-PSDB; AAV37618, AAV37619.  
 XX New isolated gene associated with glaucoma - used to develop products to  
 PT determine whether a subject has, or is at risk of, developing glaucoma,  
 PT and for treating or preventing glaucoma.  
 XX Disclosure; Page 92-94; 116pp; English.

XX This represents a human GLCIA protein sequence. The human GLCIA gene is  
 CC associated with juvenile open angle glaucoma (JOAG). The gene can be used  
 CC for the development of assays for identifying molecules that modulate  
 CC (agonists or antagonists) the bioactivity of a functional or mutant gene  
 CC or protein. Modulators may be an antibody, protein, peptide or  
 CC peptidomimetic or a nucleic acid, e.g. antisense sequence, ribozyme or  
 CC triple helix forming nucleic acid. These molecules can be administered to  
 CC a subject with glaucoma or at risk for developing glaucoma to prevent or  
 CC reduce the severity of the condition. Derivatives of GLCIA gene can be  
 CC used to detect lesions of the GLCIA gene which are indicative of glaucoma  
 CC or predisposition to glaucoma

XX SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 101  
 AAW70496  
 ID AAW70496 standard; protein; 504 AA.  
 XX AC AAW70496;  
 XX DT 11-JAN-1999 (first entry)  
 XX DE Trabecular meshwork induced glucocorticoid response\* protein.  
 XX KW Trabecular meshwork induced glucocorticoid response\*; TIGR\*; human;  
 XX OS glaucoma; diagnosis.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Peptide 1..20 /label= Sig\_peptide  
 FT Protein 21..504 /label= Mat\_protein  
 FT Modified-site 57..60 /note= "Asn is N-glycosylated"  
 FT Region 85..92 /note= "consensus leucine zipper unit"  
 FT Region 92..99 /note= "consensus leucine zipper unit"  
 FT Region 131..138 /note= "consensus leucine zipper unit"  
 FT Region 138..145 /note= "consensus leucine zipper unit"  
 FT Region 145..152 /note= "consensus leucine zipper unit"  
 FT Region 152..159 /note= "consensus leucine zipper unit"  
 FT Region 156..160 /note= "consensus leucine zipper unit"  
 FT Binding-site 159..166 /note= "heparin sulphate binding"  
 FT Region 231..232 /note= "consensus leucine zipper unit"  
 FT Modified-site 232..233 /note= "O-glycosylation"  
 FT Modified-site 233..234 /note= "O-glycosylation"  
 FT Domain 238..239 /note= "initiation domain"  
 FT Domain 277..279 /note= "initiation domain"  
 FT Modified-site 312..313 /note= "O-glycosylation"  
 FT Modified-site 331..332 /note= "O-glycosylation"  
 FT Domain 404..408 /note= "initiation domain"  
 FT Modified-site 460..464 /note= "O-glycosylation"  
 FT Modified-site 464..466 /note= "O-glycosylation"  
 XX WO9844108-A1.  
 PN

PD 08-OCT-1998.  
 XX  
 PF 07-APR-1997; 97WO-US005801.  
 XX  
 PR 01-APR-1997; 97WO-US005391.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Nguyen TD, Polansky JR, Huang W;  
 XX WPI; 1998-542701/46.  
 DR N-PSDB; AAV33484.  
 XX  
 XX  
 PT New protein induced in trabecular meshwork cells by glucocorticoids -  
 XX useful in the diagnosis of glaucoma and related diseases.  
 XX  
 PS Claim 1; Fig 1A-C; 53pp; English.  
 XX  
 CC This is the amino acid sequence of a new human 55 kDa protein, designated  
 CC trabecular meshwork induced glucocorticoid response\* (TIGR\*) protein,  
 CC that is highly induced by glucocorticoids in the endothelial lining of  
 CC the human trabecular meshwork (HTM). The sequence was deduced from an  
 CC isolated cDNA clone (see AAV33484). Studies of the recombinant protein  
 CC suggest (1) that the 55 kDa protein exists both in cells and in the  
 CC medium, (2) that it undergoes oligomerisation, (3) phosphorylation, (4)  
 CC glycosylation, (5) that it is susceptible to metalloprotease, (6) that it  
 CC exhibits high affinity binding to extracellular matrix and HTM cells, (7)  
 CC that it inhibits progressive inductions with time in both cell and organ  
 CC cultures, and (8) that it exhibits high expression in the HTM of  
 CC glaucomatous patients as compared to normal patients. TIGR\* cDNA, the  
 CC protein itself, molecules that bind it, and nucleic acid molecules that  
 CC encode it, provide improved methods and reagents for diagnosing glaucoma  
 CC and related disorders, such as cardiovascular and immunological diseases  
 CC that affect expression of TIGR\*. A claimed method of diagnosing glaucoma  
 CC involves determining if the amount of TIGR\* present in the HTM exceeds  
 CC the amount found in an individual not predisposed to the disease  
 XX  
 SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 RESULT 102  
 AAY07393  
 ID AAY07393 standard; protein; 504 AA.  
 XX  
 AC AAY07393;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE Human TIGR/MYOC protein.  
 XX  
 KW Trabecular meshwork induced glucocorticoid response; TIGR; MYOC; GLCIA;  
 KW locus; chromosome; detection; mutant; allele; heterozygote; mutation;  
 KW juvenile open-angle glaucoma; phenotype; homoallelic complementation;  
 KW autosomal dominant disease; homozygote; epilepsy; mental retardation;  
 KW haploinsufficiency.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9916898-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 29-SEP-1998; 98WO-CA000923.  
 XX  
 PR 30-SEP-1997; 97CA-02216997.  
 XX

PR 12-MAY-1998; 98CA-02231720.  
 XX  
 PA (UYLA-) UNIV LAVAL.  
 XX  
 PI Raymond V, Morissette J, Falardeau P, Cote G, Anctil J;  
 XX WPI; 1999-263703/22.  
 DR N-PSDB; AAX57606.  
 XX  
 XX  
 PT Molecular diagnosis of glaucomas associated with chromosome 1.  
 XX  
 PS Disclosure; Fig 1A-J; 66pp; English.  
 XX  
 CC This sequence represents the protein encoded by the trabecular meshwork  
 CC induced glucocorticoid response (TIGR) gene, also known as MYOC, which is  
 CC mapped to the GLCIA locus on chromosome 1q23-q25. The invention relates  
 CC to the detection of mutant and non-mutant alleles of the TIGR/MYOC gene.  
 CC Juvenile open-angle glaucoma, in a heterozygotic carrier of TIGR  
 CC mutations, can be treated by overexpression of mutated TIGR, which  
 CC renders the phenotype of the patient normal by homoallelic  
 CC complementation. This method of homoallelic complementation can be used  
 CC in other autosomal dominant diseases where mutant homozygotes are  
 CC phenotypically normal, e.g. a form of epilepsy and mental retardation  
 CC linked to chromosome X, which only affects women. The methods can also be  
 CC used for treating haploinsufficiency  
 XX  
 SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 RESULT 103  
 AAW73500  
 ID AAW73500 standard; protein; 504 AA.  
 XX  
 AC AAW73500;  
 XX  
 DT 26-FEB-1999 (first entry)  
 XX  
 DE Trabecular meshwork induced glucocorticoid response protein.  
 XX  
 KW TIGR protein; trabecular meshwork induced glucocorticoid response;  
 KW secretory protein; antibody; glaucoma; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5849879-A.  
 XX  
 PD 15-DEC-1998.  
 XX  
 PF 14-MAY-1996; 96US-00645900.  
 XX  
 PR 03-NOV-1994; 94US-00336235.  
 PR 20-OCT-1995; 95US-00546568.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Polansky JR, Huang W, Nguyen TD;  
 XX WPI; 1999-069807/06.  
 DR  
 XX  
 PT Antibody to trabecular meshwork protein - useful for diagnosis of  
 PT glaucoma.  
 XX  
 PS Claim 1; Col 25-28; 22pp; English.  
 XX  
 XX This sequence represents the human Trabecular meshwork induced



CC glucocorticoid response (TIGR) protein. The TIGR protein is a secretory  
 CC protein specifically bound by the antibody of the invention. The  
 CC antibody, especially in labelled form, can be used in the diagnosis of  
 CC glaucoma by detecting elevated levels of the protein in the trabecular  
 CC meshwork of the eye. Using the antibody, glaucoma is detected more  
 CC accurately

XX SQ Sequence 504 AA;

Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 104

AAW89391  
 ID AAW89391 standard; protein; 504 AA.

XX AC AAW89391;

DT 08-MAR-1999 (first entry)

XX Human trabecular meshwork induced glucocorticoid response protein.

XX Human; trabecular meshwork induced glucocorticoid response protein; TIGR;  
 KW glaucoma; primary open angle glaucoma; POAG; pigmentary glaucoma;  
 KW low tension glaucoma; intraocular pressure; steroid; corticosteroid.

XX OS Homo sapiens.

XX US5854415-A.

XX 29-DEC-1998.

XX 25-JUN-1997; 97US-00882238.

XX 03-NOV-1994; 94US-00336235.

XX 20-OCT-1995; 95US-00548568.

XX 17-MAY-1996; 96US-00649432.

XX (REGC ) UNIV CALIFORNIA.

XX Huang W, Polansky JR, Nguyen TD;

XX WPI; 1999-095006/08.

XX N-PSDB; AAW81910.

XX New isolated glaucoma-associated nucleic acids - which encode Trabecular  
 PT Meshwork Induced Glucocorticoid Response protein, used to develop  
 PT products for diagnosing glaucoma-related diseases.

XX Claim 1; Fig 1; 22pp; English.

XX The present sequence is a human secretory protein from clone II-2. The  
 CC secretory protein is designated TIGR (Trabecular Meshwork Induced  
 CC Glucocorticoid Response) protein. The protein is highly induced by  
 CC glucocorticoids in the endothelial lining cells of the human trabecular  
 CC meshwork. The TIGR polynucleotides and proteins can be used as markers  
 CC for the diagnosis of glaucoma, primary open angle glaucoma (POAG),  
 CC pigmentary glaucoma, and low tension glaucoma and their related diseases.  
 CC They can also be used to diagnose or protect an individual's sensitivity  
 CC to elevated intraocular pressure upon administration of steroids such as  
 CC glucocorticoids or corticosteroids. These products can also be used for  
 CC diagnosing other diseases or conditions that affect the expression or  
 CC activity of the protein. The products can also be formulated for  
 CC administration to patients

XX Sequence 504 AA;

Query Match 2.1%; Score 7; DB 2; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 105

AAAY49288  
 ID AAY49288 standard; protein; 504 AA.

XX AC AAY49288;

XX 07-FEB-2000 (first entry)

XX Human GLC1A polypeptide.

XX Glaucoma; PCR amplification; primary open wide angle glaucoma;  
 KW GLC1A gene; human.

XX OS Homo sapiens.

XX WO9951779-A2.

XX 14-OCT-1999.

XX 07-APR-1999; 99WO-US007671.

XX 07-APR-1998; 98US-00056285.

XX (IOWA ) UNIV IOWA RES FOUND.

XX Stone EM, Sheffield VC, Alward WLM, Fingert J;

XX WPI; 2000-022956/02.

XX N-PSDB; AA237974.

XX Determination of a predisposition to glaucoma by analysing mutations in  
 PT the GLC1A gene.

XX Disclosure; Page 117-120; 137pp; English.

XX The invention relates to a method for the determination of a  
 CC predisposition to glaucoma. The method comprises amplifying a GLC1A gene  
 CC with a primer pair selected from the sequences shown in AAZ37981-Z38008.  
 CC The primers are used to determine whether a subject has or has the  
 CC potential to develop primary open wide angle glaucoma. The present  
 CC sequence represents the human GLC1A polypeptide

XX Sequence 504 AA;

Query Match 2.1%; Score 7; DB 3; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 106

AAAY93971  
 ID AAY93971 standard; protein; 504 AA.

XX AC AAY93971;

XX 20-OCT-2000 (first entry)

XX A trabecular meshwork inducible glucocorticoid receptor protein.

XX TIGR; trabecular meshwork inducible glucocorticoid receptor; promoter;

KW glaucoma; steroid sensitivity; progressive ocular hypertension;  
 KW vision loss.

XX Homo sapiens.

XX WO200042220-A1.

PD 20-JUL-2000.

XX 11-JAN-2000; 2000WO-US000559.

XX 11-JAN-1999; 99US-00227881.

PR 07-MAY-1999; 99US-00306828.

XX (REGC ) UNIV CALIFORNIA.

XX Nguyen TD, Polansky JR, Chen P, Chen H;

XX WPI; 2000-491060/43.

DR N-PSDE; AAA57509.

XX Diagnosis, prognosis and treatment of glaucoma, based on detecting  
 PT specific polymorphisms in the promoter of the trabecular meshwork  
 PT inducible glucocorticoid receptor gene.

XX Disclosure; Fig 8; 122pp; English.

XX The present sequence represents a human TIGR (trabecular meshwork  
 CC inducible glucocorticoid receptor) protein. The primers correspond to  
 CC sequences found within the TIGR promoter and two of the exons of TIGR,  
 CC and are used in the method of the invention. The specification describes  
 CC a method for the diagnosis, prognosis and treatment of glaucoma, based on  
 CC detecting specific polymorphisms in the promoter of the TIGR gene. The  
 CC method is used for diagnosis and prognosis of glaucoma (of all types),  
 CC steroid sensitivity and progressive ocular hypertension that leads to  
 CC loss of vision. Glaucoma can be treated by administering an agent that  
 CC binds to cis-acting elements within the TIGR promoter. The TIGR promoter  
 CC (or other regulatory regions) can be used to express homologous or  
 CC heterologous genes, particularly for tissue-specific expression of  
 CC therapeutic transgenes for treating glaucoma, also to generate transgenic  
 CC animals and in screening for compounds (specific modulators) with  
 CC diagnostic or therapeutic potential. Fragments of the TIGR sequence can  
 CC be used as amplification primers or probes, e.g. for isolating related  
 CC sequences in non-human animals

XX Sequence 504 AA;

Query Match 2.18; Score 7; DB 3; Length 504;

Best Local Similarity 100.0%; Pred. No. 7.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266

Db 144 NLLRDKS 150

RESULT 107

AAG51051

ID AAG51051 standard; protein; 504 AA.

XX AAG51051;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 64756.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridization assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.  
 PD 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 29-MAR-1999; 99US-0126264P.  
 PR 25-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144336P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147316P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148585P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151085P.  
PR 27-AUG-1999; 99US-0151086P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154019P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-01551139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157153P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160800P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 504;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 272 EAFHVK 278  
Db 439 EAFHVK 445

## RESULT 108

AAE03744  
ID AAE03744 standard; protein; 504 AA.

XX AC AAE03744;

XX DT 04-SEP-2001 (first entry)

XX DE Human trabecular meshwork induced glucocorticoid response (TIGR) protein.  
XX KW Human; trabecular meshwork induced glucocorticoid response; TIGR;  
KW therapy; glucocorticoid; glaucoma; cardiovascular disorder; steroid;  
KW immunological disease; intraocular pressure; chromosome 1; chromosome 10;  
XX KW chromosome 11; chromosome 12.

OS Homo sapiens.

XX Key Location/Qualifiers  
FT Peptide 1..15  
FT Protein /label= Signal peptide  
FT Modified-site /note= "Mature human TIGR protein"  
FT /note= "Asn is N-glycosylated"

FT Region 85. .92 /note= "Leucine zipper unit"  
 FT Region 92. .99 /note= "Leucine zipper unit"  
 FT Binding-site 110. .113 /note= "Heparin sulphate binding residues"  
 FT Region 121. .128 /note= "Leucine zipper unit"  
 FT Region 128. .135 /note= "Leucine zipper unit"  
 FT Region 135. .142 /note= "Leucine zipper unit"  
 FT Region 142. .149 /note= "Leucine zipper unit"  
 FT Binding-site 146. .150 /note= "Heparin sulphate binding residues"  
 FT Region 149. .156 /note= "Leucine zipper unit"  
 FT Modified-site 221. .222 /note= "O-glycosylation site"  
 FT Modified-site 222. .223 /note= "O-glycosylation site"  
 FT Domain 223. .224 /note= "Initiation domain"  
 FT Domain 231. .232 /note= "Initiation domain"  
 FT Modified-site 270. .272 /note= "O-glycosylation site"  
 FT Modified-site 305. .306 /note= "O-glycosylation site"  
 FT Domain 324. .325 /note= "Initiation domain"  
 FT Modified-site 397. .401 /note= "O-glycosylation site"  
 FT Modified-site 453. .457 /note= "O-glycosylation site"  
 FT Modified-site 457. .459 /note= "O-glycosylation site"  
 XX  
 FN US6248867-B1.  
 XX  
 PD 19-JUN-2001.  
 XX  
 PF 20-OCT-1995; 95US-00546568.  
 XX  
 PR 03-NOV-1994; 94US-00336235.  
 XX  
 PA (NGUYEN) NGUYEN T D.  
 PA (POLA) POLANSKY J R.  
 PA (HUAN) HUANG W.  
 XX  
 PI Nguyen TD, Polansky JR, Huang W;  
 XX WPI; 2001-407325/43.  
 DR N-PSDB; AAD08141.  
 XX  
 PT Novel fusion protein, useful in glaucoma diagnosis, comprises part of a  
 PT Trabecular Meshwork Induced Glucocorticoid Response (TIGR) protein and  
 PT binds to the TIGR protein.  
 XX  
 PS Claim 5; Fig 1; 22pp; English.  
 XX

CC The present sequence is human trabecular meshwork induced glucocorticoid  
 CC response (TIGR) protein from clone II.2. TIGR gene is mapped to p36 of  
 CC chromosome 1 and to p13, q15 of chromosome 10, 11 or 12. The trabecular  
 CC network has been proposed to play an important role in the normal flow of  
 CC the aqueous and has been presumed to be the major site of outflow  
 CC resistance in glaucomatous eyes. TIGR is highly induced by  
 CC glucocorticoids in the endothelial lining of the human trabecular  
 CC meshwork and so it is useful in diagnosing glaucomas, as well as  
 CC disorders e.g. cardiovascular and immunological diseases related to  
 CC expression of TIGR. TIGR is also useful in diagnosis or prediction of an  
 CC individual's sensitivity to elevated intraocular pressure on

CC administration of steroids  
 XX  
 SQ Sequence 504 AA;

Query Match 2.1%; Score 7; DB 4; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 260 NLLRDKS 266  
 |||||  
 Db 144 NLLRDKS 150

## RESULT 109

AAU09184  
 ID AAU09184 standard; protein; 504 AA.

XX AAU09184;

XX 16-JAN-2002 (first entry)

XX Human PRO9964 polypeptide.

XX Human; PRO9964; clone DNA96973; immune-related disorder;  
 KW inflammatory disorder; infectious disorder; immunodeficiency disorder;  
 KW autoimmune disorder; renal disease; demyelinating disease; skin disease;  
 KW neoplasia; transplantation associated disease; immunosuppressive;  
 KW anti-inflammatory; antiasthmatic; antidiabetic.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Domain 37. .57 /label= Transmembrane\_domain  
 FT Modified-site 57. .62 /note= "N-myristoylation site"  
 FT Modified-site 83. .86 /note= "N-glycosylation site"  
 FT Domain 143. .164 /note= "Leucine zipper"  
 FT Domain 150. .171 /note= "Leucine zipper"  
 FT Domain 157. .178 /note= "Leucine zipper"  
 FT Modified-site 162. .168 /note= "Tyrosine kinase phosphorylation site"  
 FT Domain 164. .185 /note= "Leucine zipper"  
 FT Domain 171. .192 /note= "Leucine zipper"  
 FT Binding-site 190. .194 /note= "Glycosaminoglycan attachment site"  
 FT Modified-site 264. .267 /note= "CAMP- and cGMP-dependent protein kinase phosphorylation site"  
 FT Domain 272. .529 /label= Olfactomedin-like\_domain  
 FT Modified-site 460. .465 /note= "N-myristoylation site"  
 FT Modified-site 482. .487 /note= "N-myristoylation site"  
 FT Modified-site 484. .489 /note= "N-myristoylation site"  
 FT Peptide 528. .531 /note= "Microbodies C-terminal targeting signal"

XX WO200166740-A2.

PN 13-SEP-2001.

XX 01-MAR-2001; 2001WO-US006666.

XX 03-MAR-2000; 2000US-0187202P.

PR 21-MAR-2000; 2000US-0191015P.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 05-JUN-2000; 2000US-0209832P.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 01-DEC-2000; 2000WO-US032678.  
 XX (GETH) GENENTECH INC.  
 PA  
 XX Eaton DL, Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL;  
 PI Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI; 2001-625876/72.  
 DR N-PSDB; AAS15366.  
 XX  
 PT Nucleic acids encoding PRO polypeptides, useful for detecting and  
 PT treating immune related diseases and disorders in mammals including  
 PT autoimmune diseases, inflammatory diseases and asthma.  
 XX  
 PS Claim 10; Fig 14; 122pp; English.  
 XX  
 CC The present invention relates to the isolation of 9 novel human PRO  
 CC polypeptides and the cDNA sequences (AAS15360-AAS15368) encoding them.  
 CC The novel PRO polypeptides include PRO1356, PRO1268, PRO1884, PRO3444,  
 CC PRO3151, PRO3422, PRO9964, PRO10008 and PRO19598. The cDNA sequences  
 CC encoding these PRO polypeptides have been designated as clones DNA64886-  
 CC 1601, DNA64903-1553, DNA84318-2520, DNA87997, DNA89273, DNA92223-2567,  
 CC DNA96973, DNA101921 and DNA145887 respectively. Compositions (e.g.  
 CC vaccines) containing PRO polypeptides and methods of using these  
 CC compositions are useful in the treatment and diagnosis of immune-related  
 CC disorders. Such disorders include immune-mediated inflammatory disorders  
 CC (e.g. osteoarthritis), non-immune-mediated inflammatory disorders (e.g.  
 CC diabetes mellitus), infectious disorders (e.g. granulomatous hepatitis),  
 CC immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g.  
 CC rheumatoid arthritis), immune-related renal diseases (e.g. cirrhosis),  
 CC demyelinating diseases of the peripheral or central nervous system (e.g.  
 CC Guillain-Barre syndrome), immune-mediated skin diseases (e.g. contact  
 CC dermatitis), neoplasias and transplantation associated diseases. The  
 CC polynucleotide sequences of the invention may be used in gene therapy.  
 CC AAU09178-AAU09186 represent the novel human PRO polypeptides of the  
 CC invention  
 XX  
 SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 4; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||  
 |||||  
 RESULT 110  
 AAB48845  
 ID AAB48845 standard; protein; 504 AA.  
 XX AAB48845;  
 XX  
 XX 13-MAR-2001 (first entry)  
 DT  
 DE Human TIGR (trabecular meshwork induced glucocorticoid response) protein.  
 XX  
 XX Human; TIGR protein; trabecular meshwork induced glucocorticoid response;  
 KW glaucoma; intraocular pressure; aqueous humour outflow; diagnosis;  
 KW cardiovascular disorder; immunological disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US6150161-A.  
 PN  
 XX 21-NOV-2000.  
 PD  
 XX 24-DEC-1998; 98US-00220459.  
 PF

XX 03-NOV-1994; 94US-00336235.  
 PR 20-OCT-1995; 95US-00546568.  
 PR 17-MAY-1996; 96US-00649432.  
 PR 25-JUN-1997; 97US-00882238.  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 XX Huang W, Polansky JR, Nguyen TD;  
 PI  
 XX WPI; 2001-060016/07.  
 DR N-PSDB; AAC87528, AAC87529.  
 DR  
 XX Novel nucleic acid encoding trabecular meshwork induced glucocorticoid  
 PT response protein, useful for diagnosing glaucoma, cardiovascular and  
 PT immunological disorders.  
 XX  
 PS Example 3; Fig 1A-D; 22pp; English.  
 XX  
 CC The invention relates to cDNA (AAC87528, AAC87529) encoding human TIGR  
 CC (trabecular meshwork induced glucocorticoid response) protein (AAB48845).  
 CC The invention also relates to a transgenic cell or progeny thereof  
 CC comprising a human TIGR protein nucleic acid of the invention. TIGR  
 CC protein expression is highly induced in trabecular meshwork cells on  
 CC exposure to glucocorticoids. It is thought that TIGR protein is deposited  
 CC in the extracellular spaces of the trabecular meshwork and binds to the  
 CC surface of endothelial cells of the trabecular meshwork. This interferes  
 CC with the normal outflow of aqueous humour from the eye, leading to an  
 CC increase in intraocular pressure and resulting in glaucoma. Human TIGR  
 CC protein cDNA is useful for diagnosing glaucoma and associated disorders,  
 CC and its use provides for an improved and more accurate diagnosis of this  
 CC condition. TIGR protein cDNA may also be used in the diagnosis of other  
 CC diseases and conditions which are associated with altered expression or  
 CC activity of TIGR protein such as cardiovascular and immunological  
 CC disorders. The present sequence represents human TIGR protein  
 XX  
 SQ Sequence 504 AA;  
 Query Match 2.1%; Score 7; DB 4; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||  
 |||||  
 RESULT 111  
 ABG75692  
 ID ABG75692 standard; protein; 504 AA.  
 XX  
 AC ABG75692;  
 XX  
 DT 04-JUN-2003 (first entry)  
 DT  
 XX Human trabecular meshwork-induced glucocorticoid response (TIGR) protein.  
 XX  
 XX Human; zsig58; gonadal development; pregnancy; pubertal change;  
 KW menopause; ovarian cancer; fertility; ovarian function; pancreas;  
 KW polycystic ovarian syndrome; diabetes; eye disease; pituitary function;  
 KW osteoporosis; bone disease; wound healing; bacterial infection;  
 KW viral infection; fungal infection; analgesic; antidiabetic; vulvuary;  
 KW gynaecological; osteopathic; cytostatic; ophthalmological;  
 KW trabecular meshwork-induced glucocorticoid response; TIGR.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2002182677-A1.  
 PN  
 XX 05-DEC-2002.  
 PD  
 XX 26-FEB-2002; 2002US-00086135.  
 PF

PR 03-AUG-1998; 98US-0095199P.  
 PR 03-AUG-1999; 99US-00366448.  
 PA (ZYWO ) ZYMOGENETICS INC.  
 XX Sheppard PO, Chandrasekher YA;  
 XX WPI; 2003-328618/31.  
 XX  
 XX New pancreatic and ovarian zsig58 polypeptides useful for diagnosing or  
 PT treating disorders associated with gonadal development, pregnancy,  
 PT pubertal changes, menopause, ovarian cancer, fertility, and ovarian or  
 PT pancreatic function.  
 XX  
 PS Disclosure; Fig 1; 49pp; English.  
 XX  
 CC The invention relates to an isolated pancreatic and ovarian zsig58  
 CC polypeptide and the polynucleotide encoding it. The polypeptide,  
 CC polynucleotide and an antibody to the polypeptide are useful in  
 CC diagnosing or treating disorders associated with gonadal development,  
 CC pregnancy, pubertal changes, menopause, ovarian cancer, fertility,  
 CC ovarian function, polycystic ovarian syndrome, pancreas, diabetes, eye  
 CC disease, pituitary function, osteoporosis and other bone diseases. The  
 CC zsig58 polypeptide may also be used in promoting wound healing, in anti-  
 CC microbial applications, as a cell culture reagent in in vitro studies of  
 CC exogenous microorganism infections (e.g. bacterial, viral or fungal  
 CC infection), as an analgesic (e.g. bone pain), in identifying cells,  
 CC identifying agonists and antagonists of its activity and in preparing  
 CC antibodies. The antibody may be used for tagging cells that express  
 CC zsig58, for isolating zsig58 and for other diagnostic and therapeutic  
 CC applications. The polynucleotide is also useful in identifying a region  
 CC of the genome associated with human disease states. This sequence  
 CC represents the human trabecular meshwork-induced glucocorticoid response  
 CC (TIGR) protein, which is related to the human zsig58 polypeptide  
 XX  
 SQ Sequence 504 AA;

Query Match 2.1%; Score 7; DB 6; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 112  
 ABP72340  
 ID ABP72340 standard; protein; 504 AA.  
 XX  
 AC ABP72340;

DT 08-MAY-2003 (first entry)

DE Human myocilin.

KW Trabecular meshwork inducible glucocorticoid responsive protein; TIGR;  
 KW myocilin; MYOC; human; glaucoma; cardiovascular disease;  
 KW retinal degenerative disease; diagnosis; genetic profiling;  
 KW ophthalmological; cardiovascular-gen.

OS Homo sapiens.

PN WO2002102300-A2.

PD 27-DEC-2002.

PF 01-NOV-2001; 2001WO-US045645.

PR 22-NOV-2000; 2000US-0252420P.

PR 05-APR-2001; 2001US-0281422P.

PR 23-JUL-2001; 2001US-0306889P.

XX (KONG/) KONG T H.  
 XX Kong TH;  
 XX  
 DR WPI; 2003-210087/20.  
 DR N-PSDB; ABZ58478.  
 XX

PT Determining whether a subject has or is at risk of developing glaucoma,  
 PT retinal degenerative or cardiovascular disease by generating a  
 PT transcriptional profile in the subject, based on the expression of  
 PT myocilin.  
 XX

PS Claim 36; Fig 3; 52pp; English.

CC The present sequence is the protein sequence of human trabecular meshwork  
 CC inducible glucocorticoid responsive protein (TIGR), also known as  
 CC myocilin (MYOC). The invention is based on the finding that in addition  
 CC to full-length MYOC mRNA, shorter mRNAs of the TIGR gene are expressed in  
 CC cells of ocular and non-ocular origin of the human body. These shorter  
 CC mRNAs are differentially expressed in normal and glaucoma subjects. A  
 CC claimed method of determining whether a subject has, or is at risk of  
 CC developing, glaucoma, retinal degenerative or cardiovascular disease  
 CC comprises generating a transcriptional profile (i.e. a fingerprint) in  
 CC the subject or in a sample obtained from the subject, based on the  
 CC expression of the different MYOC mRNA species or polypeptide forms. A  
 CC difference in the profile relative to that in a normal subject indicates  
 CC that the subject has or is at risk of developing these diseases. Claims  
 CC are also included for methods of establishing a MYOC genetic population  
 CC profile in a population of individuals having glaucoma, retinal  
 CC degenerative or cardiovascular disease, and for pharmacogenomically  
 CC selecting a therapy for an individual having glaucoma, retinal  
 CC degenerative or cardiovascular disease. Fingerprinting techniques for  
 CC detection of the full-length or shorter forms of the MYOC mRNA or cDNA  
 CC include the use of microarrays and biochips. A claimed kit for  
 CC determining whether a subject has, or is likely to develop, a glaucoma,  
 CC retinal degenerative or cardiovascular disease comprises a probe  
 CC comprising about 10 consecutive amino acid residues or more from the  
 CC present sequence or naturally occurring mutants of it  
 XX

SQ Sequence 504 AA;

Query Match 2.1%; Score 7; DB 6; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 NLLRDKS 266  
 Db 144 NLLRDKS 150  
 |||||

RESULT 113  
 ABO44238  
 ID ABO44238 standard; protein; 504 AA.  
 XX  
 AC ABO44238;

DT 25-SEP-2003 (first entry)

DE Human TIGR/myocilin.

KW Human; latrophlin 3; LPH3; ophthalmological; hypotensive; gene therapy;  
 KW eye disease; primary open-angle glaucoma; ocular hypertension;  
 KW elevated intraocular pressure; TIGR; myocilin.

OS Homo sapiens.

PN US2003054347-A1.

PD 20-MAR-2003.

PF 27-APR-2001; 2001US-00844653.

PR 27-APR-2001; 2001US-00844653.  
 XX (UNMI ) UNIV MICHIGAN.  
 XX Richards JE, Rozsa FW;  
 XX WPI; 2003-521847/49.  
 DR N-PSDB; ACH03575.  
 XX  
 PT New Latrophilin (LPH) polynucleotides and polypeptides, useful for  
 PT diagnosing or treating subjects at risk for or having eye disease, e.g.  
 PT Primary Open-Angle Glaucoma, ocular hypertension, or elevated intraocular  
 PT pressure.  
 XX  
 PS Disclosure; Fig 11; 153pp; English.  
 XX  
 CC The invention describes a new composition, which comprises an isolated  
 CC Latrophilin (LPH) nucleic acid. The compositions are useful for  
 CC diagnosing or treating subjects at risk for or having eye disease, e.g.  
 CC Primary Open-Angle Glaucoma (e.g. juvenile onset or adult onset), ocular  
 CC hypertension, or elevated intraocular pressure. This is the amino acid  
 CC sequence of human TIGR/myocilin  
 XX  
 XX Sequence 504 AA;  
 SQ  
 Query Match 2.1%; Score 7; DB 6; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 DB 144 NLLRDKS 150  
 |||||  
 ADE54517  
 ID ADE54517 standard; protein; 504 AA.  
 XX  
 AC ADE54517;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX Human Protein Q99972, SEQ ID NO 320.  
 DE  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS  
 XX WO2003016475-A2.  
 PN  
 XX 27-FEB-2003.  
 PD  
 XX 14-AUG-2002; 2002WO-US025765.  
 PF  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 PR  
 XX (GEO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 XX WPI; 2003-268312/26.  
 DR GENBANK; Q99972.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PT  
 XX Claim 1; Page; 1017pp; English.  
 PS  
 XX

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 504 AA;  
 SQ

Query Match 2.1%; Score 7; DB 7; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 260 NLLRDKS 266  
 DB 144 NLLRDKS 150  
 |||||

RESULT 115  
 AAG32190  
 ID AAG32190 standard; protein; 508 AA.  
 XX  
 AC AAG32190;  
 XX  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 38785.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 XX 06-SEP-2000.  
 PD  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 16-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.



PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135622P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139892P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144844P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147433P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159631P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.  
  
 Query Match 2.1%; Score 7; DB 3; Length 508;  
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 301 LIEFLSS 307  
 DB 126 LIEFLSS 132  
  
 RESULT 116  
 AAB90586  
 ID AAB90586 standard; protein; 514 AA.  
 AC AAB90586;  
 XX  
 XX 01-JUN-2001 (first entry)  
 DT Human secreted protein, SEQ ID NO: 124.  
 DE Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW anti-inflammatory; anti-HIV; cytostatic; cardiac; vascular;  
 KW anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;  
 KW anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial;  
 KW vulnery; vaccine; gene therapy; cancer; protein coordinate data;  
 KW infection.  
 XX Homo sapiens.  
 OS  
 XX WO200121658-A1.  
 PN  
 XX 29-MAR-2001.  
 PD  
 XX 22-SEP-2000; 2000WO-US026013.  
 PF  
 XX 24-SEP-1999; 99US-0155709P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 PI Young PE, Wei P, Florence KA;  
 XX WPI; 2001-235311/24.  
 DR N-PSDB; AAF97926.  
 DR  
 XX Nucleic acids encoding 32 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 PT and diabetic retinopathy.  
 PT  
 XX Claim 1; Page 805-807; 890pp; English.  
 PS

XX The present sequence is one of 32 novel human secreted polypeptides. The  
 CC nucleic acid molecules and polypeptides may be used in the prevention,  
 CC diagnosis and treatment of diseases such as immune disorders (e.g.  
 CC multiple sclerosis, systemic lupus erythematosus and human immuno-  
 CC deficiency virus (HIV) infections), hyperproliferative disorders (e.g.  
 CC cancers and Gaucher's disease), cardiovascular diseases (e.g. Scimitar  
 CC syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis),  
 CC angiogenic disorders (e.g. corneal graft neovascularisation and diabetic  
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,  
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
 CC for promoting wound healing, regeneration and/or chemotaxis. The nucleic  
 CC acid molecules may be used to produce the secreted polypeptides. They may  
 CC also be used as DNA probes in diagnostic assays to detect and quantitate  
 CC the presence of similar nucleic acid sequences in samples. The  
 CC polypeptides may be used as antigens in the production of antibodies and  
 CC in assays to identify modulators of their expression and activity  
 XX

SQ Sequence 514 AA;  
 Query Match 2.1%; Score 7; DB 4; Length 514;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 66 VAQLAQE 72  
 DB 392 VAQLAQE 398  
  
 RESULT 117  
 ABG65472  
 ID ABG65472 standard; protein; 514 AA.  
 XX  
 AC ABG65472;  
 XX  
 XX 27-AUG-2002 (first entry)  
 DT Human albumin fusion protein #2147.  
 DE  
 XX Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 KW human serum albumin; HSA; cancer; reproductive disorder;  
 KW digestive disorder; immune disorder; endocrine disorder;  
 KW haematopoietic disorder; neural disorder; connective disorder;  
 KW cytostatic; antifertility; antinflammatory; antiulcer;  
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
 KW osteopathic; antiarthritic.  
 XX Homo sapiens.  
 OS  
 XX Synthetic.  
 XX WO200177137-A1.  
 PN  
 XX 18-OCT-2001.  
 PD  
 XX 12-APR-2001; 2001WO-US011988.  
 PF  
 XX 12-APR-2000; 2000US-0229358P.  
 PR 25-APR-2000; 2000US-0199384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Haseltine WA;  
 PI WPI; 2002-010886/01.  
 DR  
 XX New fusion protein for treating disease e.g. diabetes comprises an  
 PT albumin fused to a therapeutic protein.  
 PT  
 XX Claim 1; Page 2040-2041; 2102pp; English.  
 PS  
 XX The present invention relates to albumin fusion proteins comprising a  
 CC

CC therapeutic protein X and human albumin (HA, also known as human serum  
CC albumin, HSA). The proteins are useful for treating a disease or disorder  
CC that may be modulated by therapeutic protein X. The albumin extends the  
CC shelf-life of protein X, and may increase its biological in vitro/in vivo  
CC activity. The protein is useful for treating and diagnosing disorders  
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
CC disease, ulcerative colitis), immune disorders (e.g. acquired  
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,  
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).  
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention  
XX  
SQ Sequence 514 AA;

Query Match 2.1%; Score 7; DB 5; Length 514;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 VAQLAQE 72  
|||  
Db 392 VAQLAQE 398

RESULT 118  
ADC01554  
ID ADC01554 standard; protein; 523 AA.

XX AC ADC01554;

XX DT 04-DEC-2003 (first entry)

XX DE Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 1602.

XX KW enterohaemorrhagic; anti-bacterial.

XX OS Escherichia coli; O157:H7.

XX PN JP2002355074-A.

XX PD 10-DEC-2002.

XX PF 24-JAN-2002; 2002JP-00015959.

XX PR 24-JAN-2001; 2001JP-00112010.

XX PA (UYTS-) UNIV TSUKUBA.

XX DR WPI; 2003-451640/43.

XX PT Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule  
XX and a polypeptide and its use, a polypeptide, a vector and a host cell.

XX PS Claim 3; SEQ ID NO 1602; 2067pp; Japanese.

XX The invention relates to a novel enterohaemorrhagic Escherichia coli  
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
CC has anti-bacterial activity. The polypeptide can be used in detection  
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the  
CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present  
CC sequence represents an E. coli O157:H7-specific polypeptide of the  
CC invention.

XX SQ Sequence 523 AA;

Query Match 2.1%; Score 7; DB 7; Length 523;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 PLAKIIL 160  
|||

Db 356 PLAKIIL 362

RESULT 119

ABB58106  
ID ABB58106 standard; protein; 535 AA.

XX AC ABB58106;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 1110.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-658860/75.

XX DR N-PSDB; ABL02209.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.

XX PS Disclosure; SEQ ID NO 1110; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 535 AA;

Query Match 2.1%; Score 7; DB 4; Length 535;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 231 LLGELIL 237  
|||

Db 515 LLGELIL 521

RESULT 120

ABG16717  
ID ABG16717 standard; protein; 550 AA.

XX AC ABG16717;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #16708.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.

```

OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS80904.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 47076; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 550 AA;
XX
XX Query Match 2.1%; Score 7; DB 4; Length 550;
XX Best Local Similarity 100.0%; Pred. No. 8.1e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 73 LYSSGLL 79
XX 295 LYSSGLL 301
XX
XX RESULT 121
XX AAB85251
XX ID AAB85251 standard; protein; 551 AA.
XX
XX AC AAB85251;
XX
XX DT 07-SEP-2001 (first entry)
XX
XX DE Plant thioredoxin-porphobilinogen synthase (T-PPS).
XX
XX Thioredoxin; porphobilinogen synthase; T-PPS; PPS; enzyme; herbicide;
XX delta-aminolevulinic acid; plant growth inhibitor.
XX
XX OS Lycopersicon esculentum.
XX
XX Key Location/Qualifiers
XX FT Protein 1..167
XX FT Protein /note= "thioredoxin functional fragment"
XX FT Protein 168..551
XX FT Protein /note= "PPS"
XX
XX WO200146446-A1.
XX
XX 28-JUN-2001.
XX
XX 19-DEC-2000; 2000WO-US034584.
XX
XX 22-DEC-1999; 99US-0171785P.
XX
XX (PARA-) PARADIGM GENETICS INC.
XX
XX Crawford JM, Rice J, Sevala V, Stewart S;
XX WPI; 2001-418081/44.
XX N-PSDB; AAB22901.
XX
XX Novel plant thioredoxin-porphobilinogen synthase or porphobilinogen
XX synthase polypeptides, useful for identifying compounds for use as
XX herbicides by inhibiting enzymatic activity of the polypeptides.
XX
XX Disclosure; Page 16-17; 25pp; English.
XX
XX The invention provides novel DNA sequences encoding enzymes such as plant
XX thioredoxin-porphobilinogen synthase (T-PPS) and plant porphobilinogen
XX synthase (PPS). Methods of determining the enzymatic activity of T-PPS or
XX PPS or its functional fragment are provided that involves contacting
XX delta-aminolevulinic acid with the protein or its functional fragment and
XX measuring the amount of porphobilinogen formed from it. Compounds which
XX can modify the enzymatic activity T-PPS or PPS can also be identified
XX similarly, which are useful for inhibiting plant growth by inhibiting
XX enzymatic activity of T-PPS or PPS or its functional fragment. The
XX compounds thus identified are useful as herbicides. The present sequence
XX represents the fusion protein T-PPS
XX
XX SQ Sequence 551 AA;
XX
XX Query Match 2.1%; Score 7; DB 4; Length 551;
XX Best Local Similarity 100.0%; Pred. No. 8.1e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 261 LLRKSP 267
XX 483 LLRKSP 489
XX
XX RESULT 122
XX AAB85252
XX ID AAB85252 standard; protein; 551 AA.
XX
XX AC AAB85252;
XX
XX DT 07-SEP-2001 (first entry)
XX
XX DE Plant thioredoxin-porphobilinogen synthase (T-PPS).
XX
XX Thioredoxin; porphobilinogen synthase; T-PPS; PPS; enzyme; herbicide;
XX delta-aminolevulinic acid; plant growth inhibitor.
XX
XX OS Lycopersicon esculentum.
XX
XX Key Location/Qualifiers
XX FT Protein 1..167
XX FT Protein /note= "thioredoxin functional fragment"
XX FT Protein 168..551
XX FT Protein /note= "PPS"
XX
XX WO200146446-A1.
XX

```



[illegible]

PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0132407P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149910P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156556P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.



PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160880P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 2.1%; Score 7; DB 3; Length 567;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 292 EILLKXQ 298  
Db 25 EILLKXQ 31  
|||||  
|  
  
RESULT 125  
AAG32074  
ID AAG32074 standard; protein; 576 AA.  
XX AC AAG32074;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38626.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 08-APR-1999; 99US-0128234P.  
PR 16-APR-1999; 99US-0128714P.  
PR 19-APR-1999; 99US-0129845P.  
PR 21-APR-1999; 99US-0130077P.  
PR 23-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.

```
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-01493368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149823P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151068P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158039P.
PR 08-OCT-1999; 99US-0158222P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 576;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKQ 298
DB 34 EILLKQ 40
|||||

RESULT 126
AAG39771
ID AAG39771 standard; protein; 614 AA.
AC AAG39771;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49260.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
```

PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139879P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 18-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147322P.  
PR 05-AUG-1999; 99US-0147606P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147316P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148694P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149910P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 614;  
Best Local Similarity 100.0%; Pred. No. 8.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGLL 79  
Db 261 LYSSGLL 267

RESULT 127

AAAG32189  
ID AAG32189 standard; protein; 622 AA.

XX XX

AC AAG32189;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38784.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136332P.

XX 28-MAY-1999; 99US-0136782P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 22-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155653P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158023P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159322P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159638P.  
 PR 21-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.1%; Score 7; DB 3; Length 622;  
 Best Local Similarity 100.0%; Pred. No. 9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 301 LIEFLSS 307  
 Db 240 LIEFLSS 246  
 RESULT 128  
 AAY53000  
 ID AAY53000 standard; protein; 647 AA.  
 XX  
 AC AAY53000;  
 XX  
 DT 29-FEB-2000 (first entry)  
 XX  
 DE Human secreted protein clone dn721\_8 protein sequence SEQ ID NO:6.  
 XX  
 KW Human; secreted protein; nutritional; cytokine; cell proliferation;  
 KW differentiation; immune stimulating; vaccine; suppression;  
 KW haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic;  
 KW chemokinetic; haemostatic; thrombolytic; receptor; ligand;  
 KW anti-inflammatory; cadherin; tumour invasion suppressor;  
 KW tumour inhibition; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9957132-A1.  
 PD 11-NOV-1999.  
 XX  
 PF 07-MAY-1999; 99WO-US009970.  
 XX  
 PR 07-MAY-1999; 98US-0084564P.  
 PR 02-JUN-1998; 98US-0087645P.  
 PR 22-JUL-1998; 98US-0093712P.  
 PR 31-JUL-1998; 98US-0094935P.  
 PR 10-AUG-1998; 98US-0095880P.  
 PR 11-AUG-1998; 98US-0096068P.  
 PR 06-MAY-1999; 99US-00306111.  
 XX  
 PA (GEMY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Steininger R, Bowman MR;  
 PI DiBasilio-Smith E, Widom A;  
 XX  
 DR WPI; 2000-052937/04.  
 DR N-PSDB; AAZ33318.  
 XX  
 PT New polynucleotides encoding secreted human proteins, derived from adult  
 placenta, adult retina, fetal brain, fetal.  
 XX  
 PS Claim 15; Page 365-367; 492pp; English.  
 XX  
 CC The present invention describes new human secreted proteins which were  
 isolated from adult placenta, adult retina, foetal brain, foetal kidney,  
 adult blood, adult brain, adult thyroid, adult bladder, adult neural  
 tissue, adult testes, and adult lymph node cDNA libraries. The human  
 secreted proteins, and the polynucleotides encoding them, are predicted  
 to have biological activities which would make them suitable for  
 treating, preventing or ameliorating medical conditions in humans and  
 animals. Suggested activities include nutritional activity, cytokine and  
 cell proliferation/differentiation activity, immune stimulating (e.g. as  
 vaccines) or suppressing activity, haematopoiesis regulating activity,  
 tissue growth activity, activin/inhibin activity.  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity. The  
 CC polynucleotides are also stated to be useful for gene therapy. AAZ33316  
 CC to AAZ33373 encode human secreted proteins, and AAY52998 to AAY53060  
 CC represent human secreted proteins, given in the present invention  
 XX  
 SQ Sequence 647 AA;

Query Match 2.1%; Score 7; DB 3; Length 647;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EVVSKSL 46  
 |||||  
 Db 585 EVVSKSL 591

RESULT 129  
 ID AAO23969 standard; protein; 658 AA.  
 AC AAO23969;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human optic atrophy 1 protein.  
 XX  
 KW Anorectic; antiinflammatory; cardiant; hypotensive; antidiabetic; OPAL;  
 KW neuroprotective; pharmaceutical composition; body-weight regulation;  
 KW thermogenesis; metabolic; obesity; Syndrome X; insulin-resistance;  
 KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstone;  
 KW pancreatic dysfunction; arteriosclerosis; coronary heart disease;  
 KW hypercholesterolaemia; dyslipidaemia; osteoarthritis; ROS defence;  
 KW reactive oxygen species; neurodegenerative; mitochondrial; gene therapy;  
 KW human; optic atrophy 1.  
 XX  
 OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT Misc-difference 633..634  
 FT /note= "Encoded by TCCT"  
 XX  
 XX WO2003061681-A2.  
 PD 31-JUL-2003.  
 XX  
 XX 24-JAN-2003; 2003WO-EP000738.  
 XX  
 PR 25-JAN-2002; 2002EP-00001806.  
 PR 14-FEB-2002; 2002EP-00003473.  
 PR 28-FEB-2002; 2002EP-00004687.  
 PR 25-APR-2002; 2002EP-00009475.  
 PR 18-JUN-2002; 2002EP-00013329.  
 PR 30-DEC-2002; 2002EP-00029081.  
 XX  
 PA (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.  
 XX  
 PI Steuernagel A, Molitor A, Eulenberg K, Broenner G;  
 XX  
 DR WPI; 2003-627418/59.  
 DR N-PSDB; AAL57523.  
 XX  
 PT New pharmaceutical composition, useful for the manufacture of an agent  
 PT for diagnosing, treating or preventing disorders related to body-weight  
 PT regulation and thermogenesis, e.g., metabolic diseases such as obesity.  
 XX  
 PS Claim 3; Fig 2C; 144pp; English.  
 XX

CC The invention relates to a novel pharmaceutical composition comprising a  
 CC nucleic acid molecule or polypeptide which is a human homologue of a  
 CC Drosophila melanogaster polypeptide or polynucleotide. The composition of  
 CC the invention may be utilised during the diagnosis, study, prevention and  
 CC treatment of diseases related to body-weight regulation and thermogenesis  
 CC including metabolic disorders such as obesity, Syndrome X and insulin-  
 CC resistance syndrome and eating disorders e.g. cachexia, diabetes  
 CC mellitus, hypertension, pancreatic dysfunctions, arteriosclerosis,  
 CC coronary heart disease, hypercholesterolaemia, dyslipidaemia,  
 CC osteoarthritis and gallstones. Furthermore, disorders related to reactive  
 CC oxygen species (ROS) defence may be addressed by the invention including  
 CC neurodegenerative disorders or mitochondrial disorders. Finally, the

CC composition of the invention may be useful in gene therapy. The current  
 CC sequence is that of the human optic atrophy 1 (OPAL) protein of the  
 CC invention  
 XX  
 SQ Sequence 658 AA;  
 Query Match 2.1%; Score 7; DB 6; Length 658;  
 Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELLIL 237  
 |||||  
 Db 210 LLGELLIL 216

RESULT 130  
 ABB71254  
 ID ABB71254 standard; protein; 685 AA.  
 XX  
 AC ABB71254;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 40554.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL15357.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 40554; 21pp + Sequence Listing; English.  
 XX

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABLU16176-ABLU30511), expressed DNA  
 CC sequences (ABLU01840-ABLU16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 685 AA;  
 Query Match 2.1%; Score 7; DB 4; Length 685;  
 Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ISKPENL 255  
 |||||  
 Db 11 ISKPENL 17

[illegible]



PA (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS90116.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 56288; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 760 AA;  
 SQ  
 Query Match 2.1%; Score 7; DB 4; Length 760;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 109 IGRSPT 115  
 Db 236 IGRSPT 242  
 |||||  
 RESULT 134  
 ID ABG91571  
 XX ABG91571 standard; protein; 831 AA.  
 AC ABG91571;  
 XX  
 XX 18-NOV-2002 (first entry)  
 DE  
 XX Purine/pyrimidine triphosphate type nucleotidyltransferase #156.  
 DE  
 XX Nucleotidyltransferase; enzyme; active site engineering;  
 KW alpha-D-glucopyranosyl phosphate thymidyltransferase; Ep;  
 KW substrate specificity; nucleotide sugar;  
 KW glycosylated bioactive natural product.  
 XX  
 XX Aquifex aeolicus.  
 OS  
 XX WO200248331-A2.  
 PN  
 XX 20-JUN-2002.  
 PD  
 XX 13-DEC-2001; 2001WO-US047953.  
 PF  
 XX 13-DEC-2000; 2000US-0254927P.  
 PR  
 XX (SLOK ) SLOAN KETTERING INST CANCER RES.  
 PA

XX  
 PI Thorson JS, Nikilov DB;  
 XX WPI; 2002-608282/65.  
 DR  
 XX Nucleotidyltransferase mutated at one or more amino acids, useful in  
 PT the synthesis of nucleotide sugars.  
 PT  
 XX Claim 3; Page; 182pp; English.  
 PS  
 XX The invention relates to a Nucleotidyltransferase mutated at one or  
 CC more amino acids selected from V173, G147, W224, N112, G175, D111, E162,  
 CC T201, I200, E199, R195, L89, L89T, L109, Y146 or Y177 (with reference to  
 CC the Salmonella enterica rmlA-encoded alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase, Ep, enzyme appearing as ABG91798). The mutations  
 CC alter the substrate specificity of the enzymes. The mutants and methods  
 CC involving them are used in the synthesis of nucleotide sugars for  
 CC altering nucleotidyltransferase substrate specificity. The  
 CC nucleotidyltransferase exhibits different substrate specificity for  
 CC GTP, CTP, TTP, UTP and ATP than a non-mutated nucleotidyltransferase.  
 CC The mutant may also exhibit a high degree of sequence identity to  
 CC Salmonella enterica Lr2 alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase (Ep) and can convert a wide variety of phosphates.  
 CC The mutants can be exploited in the biosynthesis of glycosylated  
 CC bioactive natural products of pharmacological use. The present sequence  
 CC is a nucleotidyltransferase exhibiting a high degree of sequence  
 CC identity to Salmonella enterica Lr2 alpha-D-glucopyranosyl phosphate  
 CC thymidyltransferase (Ep). Note: The present sequence is not displayed  
 CC in the specification but was obtained from Genbank  
 XX  
 XX Sequence 831 AA;  
 SQ  
 Query Match 2.1%; Score 7; DB 5; Length 831;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 219 SENYTK 225  
 Db 533 SENYTK 539  
 |||||  
 RESULT 135  
 ID AAE17313  
 XX AAE17313 standard; protein; 855 AA.  
 AC AAE17313;  
 XX  
 XX 18-APR-2002 (first entry)  
 DT  
 XX Human protocadherin protein, sbg419582PROTODADHERIN #2.  
 DE  
 XX Human; therapy; wound healing disorder; vaccine; cancer; infection;  
 KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;  
 KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;  
 KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;  
 KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;  
 KW depression; cardiovascular disease; myocardial infarction; renal failure;  
 KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;  
 KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;  
 KW hypersplenism; renal disease; hypoglycaemia; gastrointestinal disease;  
 KW neutropenic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;  
 KW haemostatic; vulnery; anticonvulsant; antineoplastic; neuroprotective;  
 KW nephrotropic; vasotrophic; cytotatic; cerebroprotective;  
 KW allergy; protocadherin.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200198342-A1.  
 PN  
 XX 27-DEC-2001.  
 PD  
 XX 22-JUN-2001; 2001WO-US019929.  
 PF  
 XX

PR 22-JUN-2000; 2000US-0213156P.  
 PR 22-JUN-2000; 2000US-0213161P.  
 PA (SMFK ) SMITHKLINE BEECHAM CORP.  
 PA (SMFK ) SMITHKLINE BEECHAM PLC.  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 PI Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;  
 PI Murdock PR, Smith RP, Strum JC, Xiang Z, Xie Q, Rizni SK;  
 XX  
 DR WPI: 2002-139783/18.  
 DR N-PSDB; AND27808.  
 XX  
 PT Novel secreted and membrane-associated polypeptides and polynucleotides  
 PT useful for preventing, ameliorating or correcting dysfunction or disease  
 PT including diabetes, cancer, hypertension and growth abnormalities.  
 XX  
 PS Claim 1; Page 125-127; 138pp; English.  
 XX  
 CC The invention relates to secreted and membrane-associated polypeptides  
 CC and polynucleotides. The sequences of the invention are useful in  
 CC diagnostic assays for detecting diseases associated with inappropriate  
 CC activity or levels of these polynucleotides, and in identifying their  
 CC agonists and antagonists that are potentially useful in therapy. The  
 CC sequences of the invention are useful as vaccines for inducing  
 CC immunological response. The sequences of the invention are useful for  
 CC treating cancers, infections, autoimmune disorders, haematopoietic  
 CC disorders, wound healing disorders, cholesterol ester storage disease,  
 CC inflammation, congenital muscular dystrophy, junctional epidermolysis  
 CC bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,  
 CC viral and bacterial infections, Alzheimer's disease, asthma, arthritis,  
 CC allergies, schizophrenia, sbg442445PROA-associated disorders,  
 CC septicemia, psoriasis, inflammatory bowel disease, transplant rejection,  
 CC graft versus host disease, ischaemia, stroke, acute respiratory disease,  
 CC syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis,  
 CC brain disorders including paraspranuclear palsy, myotonic dystrophy,  
 CC depression, anxiety disorders and sleep disorders, cardiovascular  
 CC diseases including congestive heart failure and myocardial infarction,  
 CC respiratory diseases including chronic obstructive pulmonary disease,  
 CC acute bronchitis and adult respiratory distress syndrome, liver disorders  
 CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral  
 CC and non-viral hepatitis, type II diabetes mellitus, renal disease  
 CC including acute and chronic renal failure, glomerulonephritis, Fanconi's  
 CC syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia  
 CC and tendinitis, gastrointestinal diseases including intestinal  
 CC obstruction and tropical sprue, spleen disorders including hypersplenism,  
 CC Hodgkin's disease and malignant lymphoma, testicular cancer, male  
 CC reproductive diseases including low testosterone and male infertility.  
 CC The present sequence is human protocadherin protein  
 XX  
 SQ Sequence 855 AA;  
 Query Match 2.1%; Score 7; DB 5; Length 855;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 75 SSGILVT 81  
 Db 77 SSGILVT 83  
 RESULT 136  
 AAEE24242  
 ID AAEE24242 standard; protein; 873 AA.  
 XX  
 AC AAEE24242;  
 XX  
 DT 23-SEP-2002 (first entry)  
 XX  
 DE Human 57779 (protocadherin) protein.  
 XX  
 KW Human; carboxypeptidase; scramblase; protocadherin protein; gene therapy;  
 KW inflammatory disorder; diabetes mellitus; arthritis; Alzheimer's disease;  
 KW

KW neurological disorder; systemic lupus erythematosus; Parkinson's disease;  
 KW cardiovascular disorder; atherosclerosis; arrhythmia; ischaemia; anaemia;  
 KW myocardial infarction; blood clotting disorder; dermatological; virucide;  
 KW endothelial cell disorder; immunosuppressive; transgenic animal; cancer;  
 KW hypertension; drug screening; vasotropic; blood vessel related disorder;  
 KW psoriasis; nootropic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 8..25  
 FT /label= Signal\_peptide  
 FT Domain 25..120  
 FT /note= "Cadherin domain"  
 FT Protein 26..873  
 FT /note= "Human mature 57779 protein"  
 FT Domain 26..676  
 FT /note= "Extracellular domain"  
 FT Modified-site 28..31  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 36..41  
 FT /note= "N-myristylation site"  
 FT Modified-site 79..84  
 FT /note= "N-myristylation site"  
 FT Modified-site 96..98  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 108..111  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Domain 117..127  
 FT /note= "Cadherin extracellular domain signature motif"  
 FT Domain 134..229  
 FT /note= "Cadherin domain"  
 FT Modified-site 147..154  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT Modified-site 152..155  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 169..172  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 180..183  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Binding-site 181..183  
 FT /note= "RGD cell attachment sequence"  
 FT Modified-site 184..189  
 FT /note= "N-myristylation site"  
 FT Modified-site 205..207  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 209..212  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 225..227  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Domain 226..236  
 FT /note= "Cadherin extracellular domain signature motif"  
 FT Domain 243..337  
 FT /note= "Cadherin domain"  
 FT Modified-site 246..249  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 261..264  
 FT /note= "N-glycosylation site"  
 FT Modified-site 263..266  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Domain 334..344  
 FT /note= "Cadherin extracellular domain signature motif"  
 FT Modified-site 335..338  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Domain 354..444  
 FT /note= "Cadherin domain"  
 FT Modified-site 374..377  
 FT /note= "Casein kinase II phosphorylation site"  
 FT Modified-site 374..376  
 FT /note= "Protein kinase C phosphorylation site"  
 FT Modified-site 420..423  
 FT /note= "N-glycosylation site"  
 FT Modified-site 435..437

FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 441. .451  
 FT FT /note= "Cadherin extracellular domain signature motif"  
 FT FT 458. .554  
 FT FT /note= "Cadherin domain"  
 FT FT 476. .479  
 FT FT /note= "Casein kinase II phosphorylation site"  
 FT FT 476. .478  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 483. .488  
 FT FT /note= "N-myristylation site"  
 FT FT 485. .488  
 FT FT /note= "N-glycosylation site"  
 FT FT 538. .543  
 FT FT /note= "N-myristylation site"  
 FT FT 546. .549  
 FT FT /note= "N-glycosylation site"  
 FT FT 548. .550  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 551. .561  
 FT FT /note= "Cadherin extracellular domain signature motif"  
 FT FT 570. .573  
 FT FT /note= "N-glycosylation site"  
 FT FT 573. .663  
 FT FT /note= "Cadherin domain"  
 FT FT 581. .584  
 FT FT /note= "Glycosaminoglycan attachment site"  
 FT FT 608. .611  
 FT FT /note= "Casein kinase II phosphorylation site"  
 FT FT 626. .628  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 629. .632  
 FT FT /note= "Casein kinase II phosphorylation site"  
 FT FT 633. .635  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 636. .639  
 FT FT /note= "Casein kinase II phosphorylation site"  
 FT FT 676. .679  
 FT FT /note= "N-glycosylation site"  
 FT FT 677. .701  
 FT FT /note= "Transmembrane domain"  
 FT FT 686. .691  
 FT FT /note= "N-myristylation site"  
 FT FT 702. .835  
 FT FT /note= "Intracellular domain"  
 FT FT 705. .713  
 FT FT /note= "Tyrosine kinase phosphorylation site"  
 FT FT 716. .722  
 FT FT /note= "Tyrosine kinase phosphorylation site"  
 FT FT 728. .730  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 735. .738  
 FT FT /note= "Casein kinase II phosphorylation site"  
 FT FT 749. .751  
 FT FT /note= "Protein kinase C phosphorylation site"  
 FT FT 777. .782  
 FT FT /note= "N-myristylation site"  
 FT FT 795. .798  
 FT FT /note= "N-glycosylation site"  
 FT FT 835. .859  
 FT FT /note= "Transmembrane domain"  
 FT FT 860. .873  
 FT FT /note= "Extracellular domain"  
 FT FT  
 PN WO200233088-A2.  
 XX  
 XX  
 PD 25-APR-2002.  
 XX  
 XX  
 XX 22-OCT-2001; 2001WO-US046717.  
 XX  
 XX 20-OCT-2000; 2000US-0241989P.  
 PR 20-OCT-2000; 2000US-0242324P.  
 PR 23-OCT-2000; 2000US-0242519P.

XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX Meyers RA, Curtis RAJ, Kapeller-Libermann R;  
 XX WPI; 2002-471400/50.  
 DR N-PSDB; AAD39055.  
 XX  
 XX New human carboxypeptidase, scramblase, and protocadherin protein and  
 PT polynucleotides for identifying modulators for use in diagnosing,  
 PT treating diabetes mellitus, atherosclerosis, cancer, Alzheimer's disease.  
 XX  
 PS Claim 5; Page 144; 170pp; English.  
 XX  
 CC The invention relates to isolated human carboxypeptidase, scramblase and  
 CC protocadherin protein. The nucleic acids encoding the polypeptides are  
 CC designated as 23566, 33489 and 57779 respectively. Examples of the  
 CC conditions related to aberrant activity or expression of the invention  
 CC include, inflammatory disorders such as diabetes mellitus, arthritis,  
 CC systemic lupus erythematosus; neurological disorders such as Alzheimer's  
 CC disease, Parkinson's disease; cardiovascular disorders (e.g. disorders of  
 CC the heart and/or blood vessels) such as atherosclerosis, arrhythmia,  
 CC ischaemic disease, myocardial infarction; blood clotting disorders e.g.  
 CC thrombocytopenia; or cellular proliferation or differentiation disorders  
 CC e.g. cancer; and other disorders include endothelial cell disorder e.g.  
 CC psoriasis, blood vessel related disorders e.g. hypertension and disorders  
 CC involving red blood cells e.g. anaemia; and viral diseases e.g. hepatitis  
 CC B virus. The invention is useful for producing antibodies, in drug  
 CC screening assays, in competition binding assays to discover compounds  
 CC that interact with the protein, in pharmacogenomic analysis and for  
 CC monitoring therapeutic effects during clinical trials and treatment. The  
 CC antibody is useful to assess abnormal tissue distribution or abnormal  
 CC expression during development, to identify protein turnover, to assess  
 CC normal and aberrant subcellular localisation of cells in various tissues  
 CC in an organism, to diagnostically monitor protein levels in tissue in  
 CC pharmacogenomic analysis, for tissue typing, forensic identification,  
 CC inhibiting protein function and to block ligand binding. The invention is  
 CC useful in gene therapy, for expressing antigenic peptides, as probes for  
 CC determining the chromosomal positions of the polynucleotides, for  
 CC designing ribozymes, constructing host cells, transgenic animals. The  
 CC present sequence is human 57779 protein  
 XX  
 SQ Sequence 873 AA;  
 Query Match 2.1%; Score 7; DB 5; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 75 SSGLLVT 81  
 DB 77 SSGLLVT 83  
 |||||  
 |||||  
 RESULT 137  
 AAEL17312  
 ID AAEL17312 standard; protein; 888 AA.  
 XX  
 AC AAEL17312;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX  
 DE Human protocadherin protein, sbg419582PROTOCOLADHERIN #1.  
 XX  
 KW Human; therapy; wound healing disorder; vaccine; cancer; infection;  
 KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;  
 KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;  
 KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;  
 KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;  
 KW depression; cardiovascular disease; myocardial infarction; renal failure;  
 KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;  
 KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;  
 KW hypersplenism; renal disease; hypoglycaemia; gastrointestinal disease;  
 KW neutrotropic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;

KW haemostatic; vulnerary; anticonvulsant; antirheumatic; neuroprotective;  
 KW nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;  
 KW allergy; protocadherin.  
 XX Homo sapiens.  
 XX MO200198342-A1.  
 XX 27-DEC-2001.  
 XX 22-JUN-2001; 2001WO-US019929.  
 XX 22-JUN-2000; 2000US-0213156P.  
 XX 22-JUN-2000; 2000US-0213161P.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;  
 PI Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;  
 XX WPI: 2002-139783/18.  
 DR N-PSDB; AAD27807.  
 XX Novel secreted and membrane-associated polypeptides and polynucleotides  
 PT useful for preventing, ameliorating or correcting dysfunction or disease  
 PT including diabetes, cancer, hypertension and growth abnormalities.  
 XX Claim 1; Page 123-125; 138pp; English.  
 XX The invention relates to secreted and membrane-associated polypeptides  
 CC and polynucleotides. The sequences of the invention are useful in  
 CC diagnostic assays for detecting diseases associated with inappropriate  
 CC activity or levels of these polynucleotides, and in identifying their  
 CC agonists and antagonists that are potentially useful in therapy. The  
 CC sequences of the invention are useful as vaccines for inducing  
 CC immunological response. The sequences of the invention are useful for  
 CC treating cancers, infections, autoimmune disorders, haematopoietic  
 CC disorders, wound healing disorders, cholesterol ester storage disease,  
 CC inflammation, congenital muscular dystrophy, junctional epidermolysis  
 CC bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,  
 CC viral and bacterial infections, Alzheimer's disease, asthma, arthritis,  
 CC allergies, schizophrenia, Sbg42445PROA-associated disorders,  
 CC septicemia, psoriasis, inflammatory bowel disease, transplant rejection,  
 CC graft versus host disease, ischaemia, stroke, acute respiratory disease  
 CC syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis,  
 CC brain disorders including parasupranuclear palsy, myotonic dystrophy,  
 CC depression, anxiety disorders and sleep disorders, cardiovascular  
 CC diseases including congestive heart failure and myocardial infarction,  
 CC respiratory diseases including chronic obstructive pulmonary disease,  
 CC acute bronchitis and adult respiratory distress syndrome, liver disorders  
 CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral  
 CC and non-viral hepatitis, type II diabetes mellitus, renal disease  
 CC including acute and chronic renal failure, Glomerulonephritis, Fanconi's  
 CC syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia  
 CC and tendinitis, gastrointestinal diseases including intestinal  
 CC obstruction and tropical sprue, spleen disorders including hypersplenism,  
 CC Hodgkin's disease and malignant lymphoma, testicular cancer, male  
 CC reproductive diseases including low testosterone and male infertility.  
 CC The present sequence is human protocadherin protein  
 XX Sequence 888 AA;

Query Match 2.1%; Score 7; DB 5; Length 888;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 75 SSGLLVY 81  
 | | | | |  
 Db 77 SSGLLVY 83

## RESULT 138

ABU42062  
 ID ABU42062 standard; protein; 891 AA.  
 XX  
 AC ABU42062;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #27589.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Pseudomonas syringae.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI: 2003-029926/02.  
 DR N-PSDB; ACA45932.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 69986; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 891 AA;

Query Match 2.1%; Score 7; DB 6; Length 891;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQAQEL 73  
 |||||  
 Db 639 AQAQEL 645

RESULT 139  
 ABG22638  
 ID ABG22638 standard; protein; 896 AA.

XX AC ABG22638;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #22629.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS86825.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 52997; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 896 AA;

Query Match 2.1%; Score 7; DB 4; Length 896;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RRQIGTR 112  
 |||||  
 Db 566 RRQIGTR 572

RESULT 140

AAU76150  
 ID AAU76150 standard; protein; 922 AA.

XX AC AAU76150;

XX DT 21-MAY-2002 (first entry)

XX DE Rice lipoxigenase protein.

XX KW Rice; lipoxigenase; RCI-1; transgenic; plant; plant antifungal;  
 XX KW rice chemically induced cDNA; promoter; transit peptide; plastid;  
 XX KW fungal mycotoxin inhibitor; plant breeding; enzyme.

XX OS Oryza sativa.

XX PH Key Location/Qualifiers

XX FT Peptide 1..37

XX FT /label= Transit\_peptide

XX FT /note= "This sequence is specifically claimed in claim  
 17"

XX FT Protein 38..922

XX FT /label= Mature\_lipoxigenase

XX PN WO200206490-A1.

XX PD 24-JAN-2002.

XX PF 12-JUL-2001; 2001WO-EP008085.

XX PR 13-JUL-2000; 2000GB-00017275.

XX PR 15-SEP-2000; 2000GB-00022739.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PA (UYZU-) UNIV ZUERICH.

XX PI Dudler R, Schaffrath U, Lawton KA;

XX DR WPI; 2002-188550/24.

XX DR N-PSDB; ABK15653.

XX PT Novel isolated nucleic acid encoding a promoter which is capable of  
 PT driving chemically inducible but not wound- or pathogen-inducible  
 PT expression of an associated nucleotide sequence.

XX PS Claim 33; Page 62-65; 88pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (a promoter of  
 CC rice chemically induced cDNA (RCI-1), which encodes a lipoxigenase)  
 CC capable of driving chemically-inducible but not wound- or pathogen-  
 CC inducible expression of an associated nucleotide sequence. Also included  
 CC are the RCI-1 cDNA, its encoded protein, a 4.5kb genomic clone for the  
 CC lipoxigenase gene, promoter fragments, the lipoxigenase transit peptide  
 CC which directs expressed proteins to the plastid, a vector comprising the  
 CC promoter or fragments and a transgenic plant comprising the vector. The  
 CC promoter or fragments are useful for expressing a nucleotide sequence of  
 CC interest. The transit peptide is useful for targeting an associated  
 CC protein of interest to plastids. A nucleic acid which expresses  
 CC polypeptide having lipoxigenase activity is useful for inhibiting fungal  
 CC mycotoxins when transformed into a plant. The lipoxigenase is useful for  
 CC inhibiting fungal mycotoxins. The promoter is useful for regulating  
 CC transcription of a chemically inducible but not wound or pathogen  
 CC inducible gene, which involves applying a chemical regulator to a plant  
 CC or seed containing a chemically regulatable nucleotide sequence.  
 CC Transgenic plants as described above are useful for breeding improved

CC plant lines that for example increase the effectiveness of conventional  
 CC methods such as herbicide or pesticide treatment or allow to dispense  
 CC with the methods due to their modified genetic properties. New crops with  
 CC improved stress tolerance can be obtained that, due to their optimised  
 CC genetic equipment yield harvested product of better quality than products  
 CC that were not able to tolerate comparable adverse developmental  
 CC conditions. The present sequence represents rice lipoxigenase  
 XX  
 XX

SQ Sequence 922 AA;

Query Match 2.1%; Score 7; DB 5; Length 922;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLLTRHK 195

DB 64 DLLTRHK 70

RESULT 141

ABB09658  
 ID ABB09658 standard; protein; 961 AA.

XX AC ABB09658;

XX DT 29-MAY-2002 (first entry)

XX DE Amino acid sequence of mitochondrial protein MSP1 isoform MSP1-F.

XX KW Human; mitochondrial protein; MSP1; dynamin; GTPase; OPAL;

XX KW mitochondrial function; cell division; dominant optical atrophy 1;

XX KW 7 MSP1-X; neurodegenerative disease; optical neuropathy;

XX KW antiproliferative; cancer; gene therapy.

XX OS Homo sapiens.

XX FN WO200200878-A2.

XX PD 03-JAN-2002.

XX PF 25-JUN-2001; 2001WO-FR001999.

XX PR 26-JUN-2000; 2000FR-00008140.

XX PA (UYVO-) UNIV TOULOUSE SABATIER PAUL.

XX PA (INRM) INST NAT SANTE & RECH MEDICALE.

XX PI Lenaers G, Ducommun B, Hamel C, Delettre C, Belenguer P;

XX DR WPI; 2002-139920/18.

XX DR N-PSDB; ABL41815.

XX PT New human mitochondrial polypeptide MSP1 and its isoforms, useful for  
 PT identifying agents for treating neurodegeneration, also related nucleic  
 PT acid.

XX PS Claim 2; Page 56-60; 76pp; French.

XX CC The present sequence represents an isoform of a human mitochondrial  
 CC polypeptide designated MSP1. MSP1 belongs to the dynamin family, and is a  
 CC GTPase which is implicated in mitochondrial function and cell division.  
 CC MSP1 mutants are associated with dominant optical atrophy 1 (OPAL). Seven  
 CC isoforms of human MSP1 also exist. MSP1 and its isoforms, are used to  
 CC screen for specific modulators, potentially useful for treating disorders  
 CC of mitochondrial function and cell division, especially neurodegenerative  
 CC diseases, particularly optical neuropathy and specifically OPAL. MSP1 and  
 CC its isoforms may also be used for development of antiproliferative  
 CC treatments, e.g. for cancer. Vectors that contain the nucleic acid  
 CC encoding MSP1 are useful for gene therapy

XX SQ Sequence 961 AA;

Query Match

2.1%; Score 7; DB 5; Length 961;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LLGELIL 237

DB 174 LLGELIL 180

RESULT 142

AAR44194  
 ID AAR44194 standard; protein; 962 AA.

XX AC AAR44194;

XX DT 25-MAR-2003 (revised)

XX DT 11-SEP-1995 (first entry)

XX DE Rat NMDA receptor subunit, NR2C.

XX KW Polymerase chain reaction; NMDA; N-methyl D-aspartic acid.

XX OS Rattus rattus.

XX PN DE4216321-A1.

XX PD 18-NOV-1993.

XX PF 16-MAY-1992; 92DE-04216321.

XX PR 16-MAY-1992; 92DE-04216321.

XX PA (BADI) BASF AG.

XX PI Bach A, Herb A, Monyer H, Seeburg PH;

XX DR WPI; 1993-369850/47.

XX DR N-PSDB; AAO50997.

XX PT DNA coding for N-methyl-D-aspartic acid receptor subunits - useful for  
 PT identifying N-methyl-D-aspartic acid receptor ligands.

XX PS Claim 1; Fig 3; 25pp; German.

XX CC The sequences given in AAR44192-94 represent the rat N-methyl D-aspartic  
 CC acid (NMDA) receptor subunits, NR2A, NR2B and NR2C, respectively. The  
 CC cDNA encoding these proteins were isolated using the primers given in  
 CC AAO50993-94 which were derived from a strongly conserved peptide region  
 CC within the glutamate receptor subunit transmembrane domain III. (Updated  
 CC on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 962 AA;

Query Match 2.1%; Score 7; DB 2; Length 962;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEVAQAL 69

DB 101 TEVAQAL 107

RESULT 143

ABB09659  
 ID ABB09659 standard; protein; 979 AA.

XX AC ABB09659;

XX DT 29-MAY-2002 (first entry)

XX DE Amino acid sequence of mitochondrial protein MSP1 isoform MSP1-G.

XX KW Human; mitochondrial protein; MSP1; dynamin; GTPase; OPAL;

XX KW mitochondrial function; cell division; dominant optical atrophy 1;

KW 7 MSP1-X; neurodegenerative disease; optical neuropathy;  
 KW antiproliferative; cancer; gene therapy.

OS Homo sapiens.

PN WO200200878-A2.

XX 03-JAN-2002.

XX 25-JUN-2001; 2001WO-FR001999.

XX 26-JUN-2000; 2000FR-00008140.

XX (UYTO-) UNIV TOULOUSE SABATIER PAUL.  
 PA (INRM) INST NAT SANTE & RECH MEDICALE.

PI Lenaers G, Ducommun B, Hamel C, Delettre C, Belenguer P;

XX WPI; 2002-139920/18.

DR N-PSDB; ABL41816.

XX New human mitochondrial polypeptide MSP1 and its isoforms, useful for  
 PT identifying agents for treating neurodegeneration, also related nucleic  
 PT acid.

XX Claim 2; Page 60-64; 76pp; French.

XX The present sequence represents an isoform of a human mitochondrial  
 CC polypeptide designated MSP1. MSP1 belongs to the dynamin family, and is a  
 CC GTPase which is implicated in mitochondrial function and cell division.  
 CC MSP1 mutants are associated with dominant optical atrophy 1 (OPA1). Seven  
 CC isoforms of human MSP1 also exist. MSP1, and its isoforms, are used to  
 CC screen for specific modulators, potentially useful for treating disorders  
 CC of mitochondrial function and cell division, especially neurodegenerative  
 CC diseases, particularly optical neuropathy and specifically OPA1. MSP1 and  
 CC its isoforms may also be used for development of antiproliferative  
 CC treatments, e.g. for cancer. Vectors that contain the nucleic acid  
 CC encoding MSP1 are useful for gene therapy

XX Sequence 979 AA;

Query Match 2.1%; Score 7; DB 5; Length 979;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 231 LLGELIL 237

Db 192 LLGELIL 198

RESULT 144

ABB09657

ID ABB09657 standard; protein; 997 AA.

XX ABB09657;

XX 29-MAY-2002 (first entry)

XX Amino acid sequence of mitochondrial protein MSP1 isoform MSP1-E.

XX Human; mitochondrial protein; MSP1; dynamin; GTPase; OPA1;

XX mitochondrial function; cell division; dominant optical atrophy 1;

XX 7 MSP1-X; neurodegenerative disease; optical neuropathy;

XX antiproliferative; cancer; gene therapy.

OS Homo sapiens.

XX WO200200878-A2.

XX 03-JAN-2002.

XX 25-JUN-2001; 2001WO-FR001999.

PR 26-JUN-2000; 2000FR-00008140.

XX (UYTO-) UNIV TOULOUSE SABATIER PAUL.  
 PA (INRM) INST NAT SANTE & RECH MEDICALE.

PI Lenaers G, Ducommun B, Hamel C, Delettre C, Belenguer P;

XX WPI; 2002-139920/18.

DR N-PSDB; ABL41814.

XX New human mitochondrial polypeptide MSP1 and its isoforms, useful for  
 PT identifying agents for treating neurodegeneration, also related nucleic  
 PT acid.

XX Claim 2; Page 52-56; 76pp; French.

XX The present sequence represents an isoform of a human mitochondrial  
 CC polypeptide designated MSP1. MSP1 belongs to the dynamin family, and is a  
 CC GTPase which is implicated in mitochondrial function and cell division.  
 CC MSP1 mutants are associated with dominant optical atrophy 1 (OPA1). Seven  
 CC isoforms of human MSP1 also exist. MSP1, and its isoforms, are used to  
 CC screen for specific modulators, potentially useful for treating disorders  
 CC of mitochondrial function and cell division, especially neurodegenerative  
 CC diseases, particularly optical neuropathy and specifically OPA1. MSP1 and  
 CC its isoforms may also be used for development of antiproliferative  
 CC treatments, e.g. for cancer. Vectors that contain the nucleic acid  
 CC encoding MSP1 are useful for gene therapy

XX Sequence 997 AA;

Query Match 2.1%; Score 7; DB 5; Length 997;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 231 LLGELIL 237

Db 210 LLGELIL 216

RESULT 145

ABB09660

ID ABB09660 standard; protein; 1015 AA.

XX ABB09660;

XX 29-MAY-2002 (first entry)

XX Amino acid sequence of mitochondrial protein MSP1 isoform MSP1-H.

XX Human; mitochondrial protein; MSP1; dynamin; GTPase; OPA1;

XX mitochondrial function; cell division; dominant optical atrophy 1;

XX 7 MSP1-X; neurodegenerative disease; optical neuropathy;

XX antiproliferative; cancer; gene therapy.

OS Homo sapiens.

XX WO200200878-A2.

XX 03-JAN-2002.

XX 25-JUN-2001; 2001WO-FR001999.

XX 26-JUN-2000; 2000FR-00008140.

XX (UYTO-) UNIV TOULOUSE SABATIER PAUL.  
 PA (INRM) INST NAT SANTE & RECH MEDICALE.

PI Lenaers G, Ducommun B, Hamel C, Delettre C, Belenguer P;

XX WPI; 2002-139920/18.

DR N-PSDB; ABL41817.

XX New human mitochondrial polypeptide MSP1 and its isoforms, useful for



PT identifying agents for treating neurodegeneration, also related nucleic  
PT acid.  
PS Claim 2; Page 64-68; 76pp; French.  
XX  
CC The present sequence represents an isoform of a human mitochondrial  
CC polypeptide designated MSPI. MSPI belongs to the dynamin family, and is a  
CC GTPase which is implicated in mitochondrial function and cell division.  
CC MSPI mutants are associated with dominant optical atrophy 1 (OPA1). Seven  
CC isoforms of human MSPI also exist. MSPI, and its isoforms, are used to  
CC screen for specific modulators, potentially useful for treating disorders  
CC of mitochondrial function and cell division, especially neurodegenerative  
CC diseases, particularly optical neuropathy and specifically OPA1. MSPI and  
CC its isoforms may also be used for development of antiproliferative  
CC treatments, e.g. for cancer. Vectors that contain the nucleic acid  
CC encoding MSPI are useful for gene therapy  
XX  
SQ Sequence 1015 AA;  
Query Match 2.1%; Score 7; DB 5; Length 1015;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 231 LLGELIL 237  
Db 228 LLGELIL 234  
RESULT 146  
ABU41624  
ID ABU41624 standard; protein; 1053 AA.  
XX  
AC ABU41624;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #27151.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Pseudomonas syringae.  
XX  
FN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699F.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA45494.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 69548; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1053 AA;  
Query Match 2.1%; Score 7; DB 6; Length 1053;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 81 TLIADLQ 87  
Db 647 TLIADLQ 653  
RESULT 147  
AAG70871  
ID AAG70871 standard; protein; 1072 AA.  
XX  
AC AAG70871;  
XX  
DT 27-JUL-2001 (first entry)  
XX  
DE C albicans apoptosis associated protein #51.  
XX  
KW Yeast; fungus; apoptosis; infection; proliferative disease; vaccine;  
KW autoimmune disease; ischaemia; neurodegeneration.  
XX  
OS Candida albicans.  
XX  
FN WO200102550-A2.  
XX  
PD 11-JAN-2001.  
XX  
PF 03-JUL-2000; 2000WO-BE000077.  
XX  
PR 01-JUL-1999; 99EP-00870141.  
XX  
PA (JANC) JANSSEN PHARM NV.  
XX  
XX Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;  
PI Nelissen BJM, Reekmans RJ;  
XX  
XX WPI; 2001-367042/38.  
DR N-PSDB; AAH29907.  
XX  
XX Yeast and fungal nucleic acids encoding proteins involved in a pathway  
PT leading to programmed cell death, useful for treating proliferative  
PT disorders, yeast and fungal infections, or for preventing apoptosis in  
PT certain diseases.

XX PS Claim 24; Fig 2; 218pp; English.

XX CC The present invention provides the protein and coding sequences of a

XX CC number of apoptosis associated proteins from the yeast *Saccharomyces*

XX CC cerevisiae and the fungus *Candida albicans*. These can be used to identify

XX CC treatments for fungal and yeast infections, for proliferative diseases

XX CC and for apoptosis related diseases such as autoimmune diseases, ischaemia

XX CC and neurodegeneration. The present sequence is one of the C. albicans

XX CC proteins of the invention

XX SQ Sequence 1072 AA;

Query Match 2.1%; Score 7; DB 4; Length 1072;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 EEVSKSL 46

Db 13 EEVSKSL 19

|||||

RESULT 148

ABP58228

ID ABP58228 standard; protein; 1148 AA.

XX AC ABP58228;

XX DT 31-MAR-2003 (first entry)

XX DE Human cell adhesion and extracellular matrix protein 5.

XX KW Cell adhesion and extracellular matrix protein 5; CADECM-5; human;

XX KW anti-HIV; virucide; anti-allergic; anti-inflammatory; anti-naemic;

XX KW anti-kininase; nootropic; anticonvulsant; anti-infertility;

XX KW anti-arteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;

XX KW cytosolic; hepatotropic; dermatological; antidiabetic; nephrotropic;

XX KW angiot; thyromimetic; neuroprotective; osteopathic; antiatherosclerotic;

XX KW antiparasitic; antihelminthic; antiparasitic; uropathic; ophthalmological;

XX KW antirheumatic; haemostatic; antibacterial; protozoacide; fungicide;

XX KW synaenological; cadherin; gene therapy.

XX OS Homo sapiens.

XX PN WO200288322-A2.

XX PD 07-NOV-2002.

XX PF 01-MAY-2002; 2002WO-US013874.

XX PR 02-MAY-2001; 2001US-0288290P.

XX PR 21-MAY-2001; 2001US-0282468P.

XX PR 15-JUN-2001; 2001US-028616P.

XX PR 28-JUN-2001; 2001US-0301672P.

XX PR 04-JAN-2002; 2002US-0345008P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Yue H, Lee EA, Duggan BM, Thangavelu K, Honchell CD, Ding L;

XX PI Hillman JL, Baughn MR, Kallick DA, Lee S, Warren BA, Xu Y, Tran UK;

XX PI Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB, Gandhi AR;

XX PI Khan FA, Walhia NK, Griffin JA, Chinn AM, Elliott VS, Ramkumar J;

XX PI Arvizu CS, Forsythe IU;

XX DR WPI; 2003-167112/16.

XX DR N-PSDB; ABZ4582.

XX PT New human cell adhesion and extracellular matrix proteins, useful for

XX PT diagnosing, treating or preventing autoimmune or inflammatory disorder

XX PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,

XX PT cancer or hepatitis.

XX PS Claim 1; Page 150-153; 178pp; English.

XX CC The present sequence is the protein sequence of Incyte polypeptide

XX CC 806556CDI denoted human cell adhesion and extracellular matrix protein 5

XX CC (CADECM-5). The protein is encoded by a clone isolated from a temporal

XX CC cortex cDNA library. Homology searches indicate it to be a cadherin. The

XX CC invention provides CADECM-1 to -11 polypeptides (see ABP58224-34) and

XX CC polynucleotides (see ABZ4578-88), expression vectors, host cells,

XX CC antibodies, agonists and antagonists. These are useful for diagnosing,

XX CC treating or preventing disorders associated with aberrant expression of

XX CC CADECM, particularly cell proliferative disorders (e.g. arteriosclerosis,

XX CC atherosclerosis, cirrhosis, hepatitis, psoriasis, paroxysmal nocturnal

XX CC haemoglobinuria, polycythaemia vera, psoriasis, primary

XX CC thrombocytopenia or cancer), developmental disorders (e.g. renal

XX CC tubular acidosis, anaemia or mental retardation), neurological disorders

XX CC (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), reproductive

XX CC disorders (e.g. infertility or a disruption in the menstrual cycle), or

XX CC autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma, autoimmune

XX CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,

XX CC glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,

XX CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,

XX CC osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid

XX CC arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal,

XX CC parasitic, protozoal or helminthic infections

XX SQ Sequence 1148 AA;

Query Match 2.1%; Score 7; DB 6; Length 1148;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SSGLLVT 81

Db 77 SSGLLVT 83

|||||

RESULT 149

AAW23593

ID AAW23593 standard; protein; 1185 AA.

XX AC AAW23593;

XX DT 10-FEB-1998 (first entry)

XX DE Human LYST1 protein.

XX KW LYST1; human; lysosomal trafficking regulator; Chediak-Higashi syndrome;

XX KW CH syndrome; autoimmune disease; tumour; asthma; urticaria;

XX KW inflammatory bowel disease; psoriasis; systemic lupus erythematosus;

XX KW rheumatoid arthritis; systemic vasculitis; glomerulonephritis;

XX KW multiple sclerosis; post-angioplasty restenosis; vaccine; therapy;

XX KW diagnosis.

XX OS Homo sapiens.

XX PN WO9728262-A1.

XX PD 07-AUG-1997.

XX PF 31-JAN-1997; 97WO-US001748.

XX PR 01-FEB-1996; 96US-0011146P.

XX PR 20-DEC-1996; 96US-0033599P.

XX PR 23-DEC-1996; 96US-0034346P.

XX PA (UYFL) UNIV FLORIDA.

XX PI Kingsmore SF, Barbosa-Alleyne MDFS;

XX DR WPI; 1997-402616/37.

XX PT Mammalian lysosomal trafficking regulators LYST1, LYST2 and LYST2

XX PT - useful to diagnose Chediak-Higashi syndrome.

PS Claim 3; Page 129-133; 237pp; English.

XX This protein sequence comprises a lysosomal trafficking regulator (LYST1) polypeptide that is encoded by a gene that is mutated in several Chediak-Higashi syndrome (CHS) patients. LYST1 can be expressed in host cells using an isolated LYST1 gene for use in various pharmacological and immunological applications. LYST1 products (see AAW23596-97) of long and short isoforms of human LYST1 are also claimed. LYST1 regulates degranulation of lysosomes, late endosomes and acidic secretory granules, primarily in leukocytes. Inhibition of such degranulation using dominant negatively acting truncated LYST1 peptides may be used to treat inflammatory and autoimmune diseases e.g. asthma, urticaria, inflammatory bowel disease, systemic lupus erythematosus, rheumatoid arthritis, psoriasis, systemic vasculitis, glomerulonephritis, multiple sclerosis and post-angioplasty restenosis, while LYST1 peptides that mimic or augment LYST1 function may be used to treat neoplasia. (NB. the amino acid sequence of the LYST1 protein provided in the specification does not correspond in the C-terminal region to the translated sequence of the LYST1 gene provided (see AAT74198))

XX SQ Sequence 1185 AA;

Query Match 2.1%; Score 7; DB 2; Length 1185;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLGE 234  
|||||

Db 844 SLKLLGE 850

RESULT 150  
AAR45945

ID AAR45945 standard; protein; 1239 AA.

AC AAR45945;

XX 08-AUG-1994 (first entry)

XX Glutamic acid receptor.

DE Glutamic acid receptor; synaptic signal translation; diagnosis;  
KW brain disease; nerve signal; gene therapy; NMDA; cerebellum; ICR; mouse;  
KW nerve cell necrosis.

XX Mus musculus.

PH Key Location/Qualifiers

FT Region 361..369  
/note= "This line is omitted in the specification"

FT Region 622..630  
/note= "This line is omitted in the specification"

FT Region 883..891  
/note= "This line is omitted in the specification"

XX JP06014783-A.

XX 25-JAN-1994

XX 30-JUN-1992; 92JP-00173155.

XX 30-JUN-1992; 92JP-00173155.

XX (MITU ) MITSUBISHI KASEI CORP.

XX WPI; 1994-061478/08.

XX N-PSDB; AAQ56916.

XX New glutamic acid receptor and gene - for use in analysis of synaptic signal translation, and diagnosis of brain disease.

XX Claim 1; Page 25-31; 35pp; Japanese.

CC The sequence shows a glutamic acid receptor. The receptor is useful for the analysis of nerve signal translation; within the synapse, expression of synapse plasticity, nerve cell necrosis, brain structure and brain disease. It can also be used in gene therapy

XX SQ Sequence 1239 AA;

Query Match 2.1%; Score 7; DB 2; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEVAQL 69  
|||||

Db 101 TEVAQL 107

RESULT 151  
ABB77986

ID ABB77986 standard; protein; 1292 AA.

AC ABB77986;

XX 22-OCT-2002 (first entry)

XX Amino acid sequence of an Arabidopsis RAD50 homologue.

DE Nucleic acid integration; homologous recombination; telomeric region;  
KW RAD50.

XX Arabidopsis thaliana.

XX EP1217074-A1.

XX 26-JUN-2002.

XX 22-DEC-2000; 2000EP-00204693.

XX 22-DEC-2000; 2000EP-00204693.

XX (UYLE-) RIJUNIV LEIDEN.

XX (BINA-) STICHTING BINAIR VECTOR SYSTEM.

XX Hooykaas PUJ, Van Attikum H, Bundock P;  
WPI; 2002-550409/59.

PT Directing integration of nucleic acid of interest to a sub-telomeric region in an eukaryote with preference for non-homologous recombination, by steering an integration pathway towards homologous recombination.

XX Disclosure; Fig 5; 63pp; English.

XX The specification describes a method for directing integration of a nucleic acid of interest to a pre-determined site, where the nucleic acid has homology at or around the pre-determined site, in a eukaryote with a preference for non-homologous recombination. The method comprises steering an integration pathway towards homologous recombination. The method is useful for directing integration of a nucleic acid of interest to a subtelomeric and/or telomeric region in an eukaryote with a preference for non-homologous recombination. The nucleic acid of interest comprises an inactive gene to replace an active gene, or vice versa, is a portion of a gene delivery vehicle, confers a desired property to the eukaryotic cell, or encodes a therapeutic proteinaceous substance or a substance conferring resistance for an antibiotic substance to a cell. The method is useful for improving gene targeting efficiency. The method is useful in the replacement of an active gene by an inactive gene, for e.g. for the inactivation of genes controlling undesired side branches of metabolic pathways, to increase the quality of bulk products such as starch, or to increase the production of specific secondary metabolites or to inhibit formation of unwanted metabolites, and also to inactivate genes controlling senescence in fruits and flowers or that determine flower pigments. The method is also useful for replacing an inactive gene by an active gene. For e.g. the replacement of a defective p53 by an

CC intact p53. Many tumours acquire a mutation in p53 during their  
 CC development which renders it inactive and often correlates with a poor  
 CC response to cancer therapy. By replacing the defect p53 by an intact p53,  
 CC e.g. through gene therapy, conventional anti cancer therapy have better  
 CC changes of succeeding. The method is also useful for therapeutic  
 CC proteinaceous substance integration. A tumoricidal gene can be delivered  
 CC to a pre-determined site present only in e.g. proliferating cells, or  
 CC present only in tumour cells, e.g. to the site where a tumour antigen is  
 CC expressed form. AB877984-86 represent RAD50 homologues. RAD50 is involved  
 CC in non-homologous recombination

XX SQ Sequence 1292 AA;

Query Match 2.1%; Score 7; DB 5; Length 1292;

Best Local Similarity 100.0%; Pred. No. 1.8e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 AEIVKIL 21

|||||

Db 1058 AEIVKIL 1064

RESULT 152

ABG76501

ID ABG76501 standard; protein; 1537 AA.

XX AC ABG76501;

XX DT 05-NOV-2002 (first entry)

XX XX

DE DNA encoding protein modification and maintenance molecule #5.

XX Protein modification and maintenance molecule; gastrointestinal disorder;  
 KW dysphagia; esophageal spasm; gastritis; anorexia; nausea; hypertension;  
 KW cardiovascular disorder; atherosclerosis; vasculitis; aneurysm; allergy;  
 KW ischaemic heart disease; autoimmune disorder; inflammatory disorder;  
 KW acquired immunodeficiency syndrome; AIDS; ankylosing spondylitis; cancer;  
 KW anaemia; amyloidosis; cell proliferative; arteriosclerotic bursitis;  
 KW cirrhosis; developmental disorder; renal tubular acidosis; anaemia;  
 KW bone resorption; epilepsy; epithelial disorder; keratosis pilaris;  
 KW allergic contact dermatitis; insect bite; keloid; dermatofibroma; eczema;  
 KW neurological disorder; stroke; cerebral neoplasm; Alzheimer's disease;  
 KW Huntington's disease; dementia; reproductive disorder; infertility;  
 KW endometriosis; gynecomastia; ectopic pregnancy; gene therapy.

XX Homo sapiens.

XX OS

XX PN W0200260942-A2.

XX PD 08-AUG-2002.

XX PF 30-JAN-2002; 2002WO-US002813.

XX PR 31-JAN-2001; 2001US-0265705P.

XX PR 05-FEB-2001; 2001US-0266762P.

XX PR 16-FEB-2001; 2001US-0269561P.

XX PR 23-FEB-2001; 2001US-0271198P.

XX PR 01-MAR-2001; 2001US-0272813P.

XX PR 13-MAR-2001; 2001US-0275586P.

XX PR 23-MAR-2001; 2001US-0278505P.

XX PR 30-MAR-2001; 2001US-0280539P.

XX XX

XX PA (INCY-) INCYTE GENOMICS INC.

XX XX

XX PI Warren BA, Honchell CD, Lu Y, Walia NK, Burford N, Deleage AM;

XX PI Gandhi AR, Baughn MR, Griffin JA, Gietzen KJ, Lu DAM, Ison CH;

XX PI Ramkumar J, Tang TY, Lal PG, Borowski MR, Duggan BM, Hafalia AJA;

XX PI Arvizu C, Thangavelu K, Yao MG, Elliott VS, Ding L, Yue H, Lee S;

XX PI Swarnakar A, Tran UK, Xu Y;

XX XX

XX DR WPI; 2002-608499/65.

XX DR N-PSDB; ABS58372.

XX XX

PT New protein modification and maintenance molecules useful for treating or  
 PT preventing gastrointestinal, cardiovascular, autoimmune/inflammatory,  
 PT cell proliferative, developmental, neurological and reproductive  
 PT disorders.

XX Claim 1; Page 137-140; 172pp; English.

XX The invention describes an isolated human polypeptide (I), a naturally  
 CC occurring amino acid sequence at least 90 % identical to the protein, or  
 CC a biologically active fragment or an immunogenic fragment of the protein.  
 CC The protein modification and maintenance molecules are useful in the  
 CC diagnosis, treatment, and prevention of gastrointestinal (e.g. dysphagia,  
 CC esophageal spasm, gastritis, anorexia or nausea), cardiovascular (e.g.  
 CC atherosclerosis, hypertension, vasculitis, aneurysm, or ischaemic heart  
 CC disease), autoimmune/inflammatory (e.g. acquired immunodeficiency  
 CC syndrome (AIDS), allergies, ankylosing spondylitis, anaemia or  
 CC amyloidosis), cell proliferative (e.g. cancers, arteriosclerotic,  
 CC bursitis, or cirrhosis), developmental (e.g. renal tubular acidosis,  
 CC anaemia, bone resorption, or epilepsy), epithelial (e.g. allergic contact  
 CC dermatitis, keratosis pilaris, insect bites, keloid, dermatofibroma or  
 CC eczema), neurological (e.g. stroke, cerebral neoplasms, Alzheimer's  
 CC disease, Huntington's disease or dementia), and reproductive disorders  
 CC (e.g. infertility, endometriosis, gynecomastia or ectopic pregnancy).  
 CC These may also be used in assessing the effects of exogenous compounds on  
 CC the expression of nucleic acid and amino acid sequences of protein  
 CC modification and maintenance molecules. Polynucleotides are useful in  
 CC somatic and germline gene therapy. This is the amino acid sequence of a  
 CC protein modification and maintenance molecule described in the invention

XX SQ Sequence 1537 AA;

Query Match 2.1%; Score 7; DB 5; Length 1537;

Best Local Similarity 100.0%; Pred. No. 2.1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 VAQLAQE 72

|||||

Db 876 VAQLAQE 882

RESULT 153

AAW23595

ID AAW23595 standard; protein; 1545 AA.

XX AC AAW23595;

XX XX

XX DT 10-FEB-1998 (first entry)

XX XX

XX DE Murine lyst1 short isoform.

XX XX

XX KW lyst1; mouse; lysosomal trafficking regulator; Chediak-Higashi syndrome;

XX CH syndrome; autoimmune disease; tumour; vaccine; therapy; diagnosis.

XX XX

XX OS Mus musculus.

XX XX

XX PN W09728262-A1.

XX XX

XX PD 07-AUG-1997.

XX XX

XX PF 31-JAN-1997; 97WO-US001748.

XX XX

XX PR 01-FEB-1996; 96US-0011146P.

XX PR 20-DEC-1996; 96US-0033599P.

XX PR 23-DEC-1996; 96US-0034346P.

XX XX

XX PA (UYFL ) UNIV FLORIDA.

XX XX

XX PI Kingmore SF, Barbosa-Alleyne MDFS;

XX XX

XX DR WPI; 1997-402616/37.

XX DR N-PSDB; AAT74200.

XX XX

XX PT Mammalian lysosomal trafficking regulators LYST1, lyst1, LYST2 and lyst2

PT - useful to diagnose Chediak-Higashi syndrome.

XX Claim 3; Page 103; 237pp; English.

XX This protein comprises the shorter isoform of murine lysosomal  
 CC trafficking regulator lys1, a protein that regulates degranulation of  
 CC lysosomes, late endosomes and acidic secretory granules primarily in  
 CC leukocytes. Its sequence was deduced from a Lys1 cDNA clone (AA74200)  
 CC that had been identified from a beige (bg) mutation critical region yeast  
 CC artificial chromosome. Alternative splicing also yields a longer isoform  
 CC (see AAW23594) of Lys1. Homologous human sequences (see AAW23596 and  
 CC AAW23597) have been identified that are associated with Chediak-Higashi  
 CC syndrome (CHS). Murine Lys1 and human LYS1 polypeptides (see AAW23593-  
 CC 98) can be expressed in recombinant host cells for use in raising  
 CC antibodies, in vaccines and in the development of therapeutic approaches  
 CC to treatment of autoimmune diseases and certain types of tumours. The  
 CC existence of an animal model of CHS with a similar genetic lesion will  
 CC assist efforts to develop novel therapies for this disease

XX Sequence 1545 AA;

Query Match 2.1%; Score 7; DB 2; Length 1545;

Best Local Similarity 100.0%; Pred. No. 2.1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 228 SLKLLGE 234

Db 1204 SLKLLGE 1210

RESULT 154

AAB90551

ID AAB90551 standard; protein; 1745 AA.

XX AAB90551;

DT 01-JUN-2001 (first entry)

DE Human secreted protein, SEQ ID NO: 89.

XX Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW anti-inflammatory; anti-HIV; cytostatic; cardiant; vascular;  
 KW anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;  
 KW anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial;  
 KW vulnerary; vaccine; gene therapy; cancer; protein coordinate data;  
 KW infection.

XX Homo sapiens.

OS WO200121658-A1.

PN 29-MAR-2001.

XX 22-SEP-2000; 2000WO-US026013.

PF 24-SEP-1999; 99US-0155709P.

PR (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 PI Young PE, Wei P, Florence KA;

XX WPI; 2001-235311/24.

DR N-PSDB; AAF97891.

XX Nucleic acids encoding 32 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 PT and diabetic retinopathy.

XX Claim 11; Fig 4; 890pp; English.

XX The present sequence is one of 32 novel human secreted polypeptides. The

CC nucleic acid molecules and polypeptides may be used in the prevention,  
 CC diagnosis and treatment of diseases such as immune disorders (e.g.  
 CC multiple sclerosis, systemic lupus erythematosus and human immuno-  
 CC deficiency virus (HIV) infections), hyperproliferative disorders (e.g.  
 CC cancers and Gaucher's disease), cardiovascular diseases (e.g. Schmitz  
 CC syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis),  
 CC angiogenic disorders (e.g. corneal graft neovascularisation and diabetic  
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,  
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
 CC for promoting wound healing, regeneration and/or chemotaxis. The nucleic  
 CC acid molecules may be used to produce the secreted polypeptides. They may  
 CC also be used as DNA probes in diagnostic assays to detect and quantitate  
 CC the presence of similar nucleic acid sequences in samples. The  
 CC polypeptides may be used as antigens in the production of antibodies and  
 CC in assays to identify modulators of their expression and activity

XX Sequence 1745 AA;

Query Match 2.1%; Score 7; DB 4; Length 1745;

Best Local Similarity 100.0%; Pred. No. 2.4e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 66 VAQLAQE 72

Db 1084 VAQLAQE 1090

RESULT 155

ABG65473

ID ABG65473 standard; protein; 1745 AA.

XX ABG65473;

DT 27-AUG-2002 (first entry)

DE Human albumin fusion protein #2148.

XX Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 KW human serum albumin; HSA; cancer; reproductive disorder;  
 KW digestive disorder; immune disorder; endocrine disorder;  
 KW haematopoietic disorder; neural disorder; connective disorder;  
 KW cytostatic; antiinfertility; antiinflammatory; antiulcer;  
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
 KW osteopathic; antiarthritis.

XX Homo sapiens.

OS Synthetic.

PN WO200177137-A1.

XX 18-OCT-2001.

XX 12-APR-2001; 2001WO-US011988.

XX 12-APR-2000; 2000US-0229358P.

XX 25-APR-2000; 2000US-0199384P.

XX 21-DEC-2000; 2000US-0256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2002-010886/01.

XX New fusion protein for treating disease e.g. diabetes comprises an

PT albumin fused to a therapeutic protein.

XX Claim 1; Page 2042-2047; 2102pp; English.

XX The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA), also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or disorder

CC that may be modulated by therapeutic protein X. The albumin extends the  
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo  
 CC activity. The protein is useful for treating and diagnosing disorders  
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
 CC disease, ulcerative colitis), immune disorders (e.g. acquired  
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
 CC hematopoietic disorders, neural disorders (e.g. Alzheimer's,  
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).  
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention  
 XX  
 SQ Sequence 1745 AA;

Query Match 2.1%; Score 7; DB 5; Length 1745;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
 |||||  
 Db 1084 VAQLAQE 1090

RESULT 156  
 AAU99587  
 ID AAU99587 standard; protein; 1762 AA.  
 AC  
 XX  
 AC AAU99587;  
 XX  
 XX 21-OCT-2002 (first entry)  
 DT  
 XX  
 XX Novel human protease.

DE  
 XX Novel human protease.

XX Novel human protein; NHP; protease; drug screening; cosmetic;  
 KW clinical trial monitoring; nutraceutical; forensic biology;  
 KW gene expression; enzyme.  
 KW  
 XX Homo sapiens.

OS  
 XX W0200255714-A2.

FN  
 XX 18-JUL-2002.

PD  
 XX 11-JAN-2002; 2002WO-US000790.

PF  
 XX 12-JAN-2001; 2001US-0261694P.

PR  
 XX (LEXI-) LEXICON GENETICS INC.

PA  
 XX Yu X, Turner CH;

PI  
 XX WPI; 2002-599670/64.

DR  
 XX N-PSDB; ABK89926.

XX New human protease - like proteins and polynucleotides encoding the  
 PT proteins, useful e.g. in mapping a unique gene to a particular  
 PT chromosome, in identifying mutations associated with a particular  
 PT disease.

XX Claim 2; Page 38-42; 42pp; English.

XX The present invention relates to the isolation of a novel human protein  
 CC (NHP), and the polynucleotide sequence encoding it. The NHP shares  
 CC sequence homology to proteases. The gene encoding the protease of the  
 CC invention maps to chromosome 9. The polypeptide and polynucleotide  
 CC sequences of the invention can be used for diagnosis, drug screening,  
 CC clinical trial monitoring, treatment of biological or medical disorders,  
 CC and cosmetic or nutraceutical applications. The polynucleotide sequence  
 CC is useful in forensic biology, in identifying mutations associated with a  
 CC particular disease, in assessing gene expression patterns, and in gene  
 CC therapy. The polypeptide sequence can be used to generate antibodies  
 CC which can be used as reagents in diagnostic or screening assays. The  
 CC present sequence represents the human protease of the invention  
 XX

SQ Sequence 1762 AA;

Query Match 2.1%; Score 7; DB 5; Length 1762;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
 |||||  
 Db 1101 VAQLAQE 1107

RESULT 157  
 AAO30834  
 ID AAO30834 standard; protein; 1762 AA.  
 XX  
 AC AAO30834;  
 XX  
 XX 22-SEP-2003 (first entry)

DT  
 XX Human cell adhesion and extracellular matrix protein (CADECM)-24.  
 DE  
 XX Human: cell adhesion and extracellular matrix protein; immune disorder;  
 KW CADECM; cancer; gene therapy.

XX Homo sapiens.

XX W02003047526-A2.

XX 12-JUN-2003.

XX 26-NOV-2002; 2002WO-US038437.

XX 30-NOV-2001; 2001US-0334343P.

PR 07-DEC-2001; 2001US-0340278P.

PR 04-JAN-2002; 2002US-0345069P.

PR 25-JAN-2002; 2002US-0351352P.

PR 14-FEB-2002; 2002US-0357168P.

PR 29-MAR-2002; 2002US-0369128P.

PR 05-APR-2002; 2002US-0370802P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD;  
 PI Deleage AM, Elliott VS, Gandhi AR, Gietzen KJ, Gorvad AE;  
 PI Griffin JA, Ho A, Jin P, Kable AE, Lal PG, Lee EA, Lee S, Lee SY;  
 PI Marquis JP, Lehr-Wason PM, Ramkumar J, Richardson TW, Sprague WW;  
 PI Swarnakar A, Tang TY, Tran B, Tran UK, Chawla NK, Warren BA, Xu Y;  
 PI Yue H, Zheng W;

XX WPI; 2003-513695/48.

DR N-PSDB; AAL62036.

XX New human cell adhesion and extracellular matrix proteins (CADECM)  
 PT polypeptide, useful for preparing a composition for treating a disease  
 PT associated with decreased expression or overexpression of CADECM e.g.,  
 PT cancer.

XX Claim 1; Page 305-309; 374pp; English.

XX The invention relates to human cell adhesion and extracellular matrix  
 CC proteins (CADECM) and nucleic acid molecules encoding such proteins.  
 CC CADECM proteins are useful for preparing a composition for diagnosing or  
 CC treating a disease or condition associated with decreased expression or  
 CC overexpression of functional CADECM e.g., immune disorders or cancer. The  
 CC invention is also useful in gene therapy. The present sequence is human  
 CC CADECM protein

XX Sequence 1762 AA;

Query Match 2.1%; Score 7; DB 7; Length 1762;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
Db 1101 VAQLAQE 1107  
RESULT 158  
ABH10246  
ID ABB10246 standard; protein; 1766 AA.  
XX AC ABB10246;  
XX DT 10-JAN-2002 (first entry)  
XX DE Human cDNA SEQ ID NO: 554.  
XX KW Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW Proliferative disorder; inflammation.  
XX OS Homo sapiens.  
XX PN WO200154474-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001349.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218280P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234233P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.



```

PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX PA
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI
XX Rosen CA, Barash SC, Ruben SM;
XX DR
XX WPI; 2001-476161/51.
XX DR N-PSDB; ABA06468.
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
XX polypeptide is used in preventing, treating or ameliorating a medical
XX condition.
XX
XX Claim 11; SEQ ID NO 554; 859pp + Sequence Listing; English.
XX
XX The present invention provides human cDNAs, proteins and related genomic
XX DNAs. These can be used in the treatment of neural, immune system,
XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
XX renal and proliferative disorders and inflammation. The present sequence
XX is a protein of the invention
XX
XX Sequence 1766 AA;
XX
XX Query Match 2.1%; Score 7; DB 4; Length 1766;
XX Best Local Similarity 100.0%; Pred. No. 2.4e+03;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 66 VAQLAQE 72
XX Db 1105 VAQLAQE 1111
XX
XX RESULT 159
XX ABP66833
XX ID ABP66833 standard; protein; 1766 AA.
XX AC ABP66833;
XX XX
XX XX 09-DEC-2002 (first entry)
XX DT
XX DE Human polypeptide SEQ ID NO 554.
XX XX
XX Human; neutropic; neuroprotective; cytosolic; dermatological; virucide;
XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;
XX antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
XX antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX OS Homo sapiens.
XX XX
XX US2002090672-A1.
XX XX
XX 11-JUL-2002.
XX XX

```

```

PF 17-JAN-2001; 2001US-00764853.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 22-AUG-2000; 2000US-0225758P.
PR 30-AUG-2000; 2000US-0226868P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 20-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-681727/73.
XX N-PSDB; ABV83805.
XX
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
XX treatment of immune, hyperproliferative, renal, respiratory, and
XX cardiovascular, reproductive, endocrine, gastrointestinal and
XX neurological disorders.
XX
XX Claim 11; SEQ ID NO 554; 369pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABV83682-ABV84101) and proteins
XX (ABP66710-ABP67129) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX

```

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases such as cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1766 AA;

Query Match 2.1%; Score 7; DB 5; Length 1766;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
 DB 1105 VAQLAQE 1111  
 |||||

RESULT 160  
 ABB64225  
 ID ABB64225 standard; protein; 1838 AA.

AC ABB64225;  
 XX  
 DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 19467.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-65686Q/75.

XX N-PSDB; ABL08328.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 19467; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 1838 AA;

Query Match 2.1%; Score 7; DB 4; Length 1838;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KNPAEIV 18  
 DB 1258 KNPAEIV 1264  
 |||||

RESULT 161  
 AAB18265

ID AAB18265 standard; protein; 2013 AA.

XX AAB18265;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:122.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;  
 KW antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.

XX WO2000025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99WO-US026796.

XX 05-NOV-1998; 98US-0107131P.

XX (HOFF/) HOFFMAN S.

XX (CARU/) CARUCCI D.

XX (GARD/) GARDNER M.

XX (VENT/) VENTER J C.

XX Hoffman S, Carucci D, Gardner M, Venter JC;

XX WPI; 2000-365347/31.

XX Proteins encoded by chromosome 2 of the human malarial parasite,  
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the  
 PT diagnosis of P.falciparum infection.

XX Disclosure; Page 285-291; 577pp; English.

XX The present invention describes proteins and their fragments (I) encoded  
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.  
 CC Also described are: (i) nucleotide sequences (ii) encoding (i); and (2)  
 CC vaccines against P. falciparum infection comprising (i) or (ii). (i) and  
 CC (ii) are useful for the development of vaccines against P. falciparum  
 CC infection. (i) and polyclonal antisera or a monoclonal antibody raised to  
 CC immunogens comprising the sequences of (i), are useful in the detection  
 CC of infection with P. falciparum. Furthermore, (i) especially when they  
 CC are rifins or secreted or membrane proteins) can aid the identification  
 CC to identify drug resistance in P. falciparum infection, or they can be used  
 CC to identify drug resistance in P. falciparum. Sequencing of the  
 CC Plasmodium chromosome 2 and the subsequent identification of proteins  
 CC encoded by it will help to expand our understanding of parasite biology,  
 CC a process hampered by the complexity of the parasitic lifecycle, and  
 CC provide new targets for vaccine and drug development. Parasite resistance  
 CC to drugs and mosquito resistance to insecticides have led to a resurgence  
 CC of malaria in many parts of the world, and there is a pressing need for  
 CC vaccines and new drugs. AAY70078 to AAA70287 and AAB18144 to AAB18352  
 CC represent nucleotide and protein sequences given in the present  
 CC invention, but which are not specifically mentioned within the  
 CC specification

XX Sequence 2013 AA;

Query Match 2.1%; Score 7; DB 3; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 IFNNILR 106  
 DB 539 IFNNILR 545  
 |||||

## RESULT 162

AB014706  
 ID AB014706 standard; protein; 3064 AA.

XX AB014706;

DT 25-AUG-2003 (first entry)

DE Novel human protein #79.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;  
 KW metabolism-related disease; obesity; central nervous system disorder;  
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;  
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;  
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;  
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;  
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;  
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;  
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;  
 KW stroke; infection.

XX Homo sapiens.

XX WO2003023002-A2.

XX 20-MAR-2003.

XX 09-SEP-2002; 2002WO-US028539.

XX 07-SEP-2001; 2001US-0318120P.

XX 07-SEP-2001; 2001US-0318130P.

XX 10-SEP-2001; 2001US-0318430P.

XX 17-SEP-2001; 2001US-0322636P.

XX 17-SEP-2001; 2001US-0322781P.

XX 17-SEP-2001; 2001US-0322816P.

XX 17-SEP-2001; 2001US-0322817P.

XX 19-SEP-2001; 2001US-0323519P.

XX 20-SEP-2001; 2001US-0323631P.

XX 20-SEP-2001; 2001US-0323636P.

XX 25-SEP-2001; 2001US-0324969P.

XX 25-SEP-2001; 2001US-0325091P.

XX 26-SEP-2001; 2001US-0324990P.

XX 17-APR-2002; 2002US-0373212P.

XX 06-SEP-2002; 2002US-00236177.

XX (CURA-) CURAGEN CORP.

XX Spytke KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

XX Gerlach VL, Vernet CAM, Ellerman K, Rothenberg ME, Guo X;

XX Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

XX Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

XX Lepley DM, Edinger SR, Burgess CE;

XX WPI; 2003-313242/30.

XX N-PSDB; ACD19399.

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

XX and polynucleotides, useful in gene therapy, e.g. for treating or

XX preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

XX stroke or infections.

XX Claim 1; Page 256; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC polypeptide, nucleic acid and antibody are useful as therapeutics,  
 CC particularly in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, which includes a pathology associated  
 CC with NOVX polypeptide. The DNA encoding the protein is useful in gene  
 CC therapy for treating the disease or condition. In particular, the NOVX  
 CC polypeptide or polynucleotide is useful for treating endocrine/  
 CC metabolism-related diseases (e.g. obesity or diabetes), central nervous  
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,  
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune  
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,  
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or  
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,  
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver  
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),  
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).  
 CC These are also useful in developing powerful assay system for functional  
 CC analysis of various human disorders, as well as in diagnostic  
 CC applications, and for monitoring the effects of drugs during clinical  
 CC trials. This is the amino acid sequence of a novel human NOV protein  
 CC

SQ Sequence 3064 AA;

Query Match 2.1%; Score 7; DB 6; Length 3064;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKDN 24

DB 1182 VKILKDN 1188

|||||

## RESULT 163

AAW22017

ID AAW22017 standard; protein; 3433 AA.

XX AAW22017;

DT 03-OCT-1997 (first entry)

DE Utrophin.

KW Utrophin; minigene; Duchenne muscular dystrophy; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 238..250

XX FT /note= "residues 238-250 (all encoded by codon NNN) may  
 XX comprise the sequence DKSIIMVLTSL, absolutely conserved  
 XX in human, mouse and rat utrophins"

XX WO9723696-A1.

XX 26-JUN-1997.

XX 19-DEC-1996; 96WO-GB003156.

XX 19-DEC-1995; 95GB-00025962.

XX 26-JUL-1996; 96GB-00015797.

XX 24-OCT-1996; 96GB-00022174.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Tinsley JM, Davies KB;

XX WPI; 1997-341687/31.

XX N-PSDB; AAT74666.

XX Nucleic acid encoding utrophin, truncated forms and related vectors -

XX also transformed mammalian cells, used for alleviating symptoms of

XX muscular dystrophy.

XX Claim 5; Fig 9; 78pp; English.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Human utrophin (AAW22017) is 395 kDa protein which shows strong sequence  
 CC similarity to dystrophin and which may protect muscle from the  
 CC consequences of dystrophin loss. It comprises an actin binding domain,  
 CC rod domain and dystrophin protein complex (DPC) binding domain. A  
 CC truncated utrophin polypeptide (AAW22016) has been produced which  
 CC includes only the actin binding and DPC binding domains, but which  
 CC retains utrophin function. Its expression in an animal model  
 CC significantly decreased the severity of the dystrophic muscle phenotype,  
 CC indicating usefulness in the treatment of muscular dystrophy. Utrophin  
 CC polypeptides can also be used to screen for substances that modulate  
 CC utrophin binding to actin and/or the DPC  
 XX  
 SQ Sequence 3433 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 3433;  
 Best Local Similarity 100.0%; Pred. No. 4.5e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 VKILKDN 24  
 |||||  
 DB 1182 VKILKDN 1188  
 RESULT 164  
 AAW23594  
 ID AAW23594 standard; protein; 3788 AA.  
 XX  
 AC AAW23594;  
 XX  
 DT 10-FEB-1998 (first entry)  
 XX  
 DE Murine Lyst1 long isoform.  
 XX  
 KW Lyst1; mouse; lysosomal trafficking regulator; Chediak-Higashi syndrome;  
 KW CH syndrome; autoimmune disease; tumour; vaccine; therapy; diagnosis.  
 XX  
 OS Mus musculus.  
 XX  
 PN W09728262-A1.  
 XX  
 PD 07-AUG-1997.  
 XX  
 PF 31-JAN-1997; 97WO-US001748.  
 XX  
 PR 01-FEB-1996; 96US-0011146P.  
 PR 20-DEC-1996; 96US-0033599P.  
 PR 23-DEC-1996; 96US-0034346P.  
 XX  
 PA (UWFL) UNIV FLORIDA.  
 XX  
 PI Kingsmore SF, Barbosa-Alleyne MDFS;  
 XX  
 XX WPI; 1997-402616/37.  
 DR N-PSDB; AAT74199.  
 XX  
 PT Mammalian lysosomal trafficking regulators LYST1, LYST2 and LYST2  
 PT - useful to diagnose Chediak-Higashi syndrome.  
 XX  
 PS Claim 3; Page 101-103; 237pp; English.  
 XX  
 CC This protein comprises the longer isoform of murine lysosomal trafficking  
 CC regulator Lyst1, a protein that regulates degradation of lysosomes,  
 CC late endosomes and acidic secretory granules primarily in leukocytes. Its  
 CC sequence was deduced from a Lyst1 cDNA clone (AAT74199) that had been  
 CC identified from a beige (bg) mutation critical region yeast artificial  
 CC chromosome. Alternative splicing also yields a shorter isoform (see  
 CC AAW23595) of Lyst1. Homologous human sequences (see AAW23596 and  
 CC AAW23597) have been identified that are associated with Chediak-Higashi  
 CC syndrome (CHS). Murine Lyst1 and human LYST1 polypeptides (see AAW23593-  
 CC 98) can be expressed in recombinant host cells for use in raising  
 CC antibodies, in vaccines and in the development of therapeutic approaches  
 CC to treatment of autoimmune diseases and certain types of tumours. The

CC existence of an animal model of CHS with a similar genetic lesion will  
 CC assist efforts to develop novel therapies for this disease.  
 XX  
 SQ Sequence 3788 AA;  
 Query Match 2.1%; Score 7; DB 2; Length 3788;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 228 SLKLLGE 234  
 |||||  
 DB 1204 SLKLLGE 1210  
 RESULT 165  
 ABP96902  
 ID ABP96902 standard; peptide; 9 AA.  
 XX  
 AC ABP96902;  
 XX  
 DT 16-JUN-2003 (first entry)  
 XX  
 DE Human CASB933 immunogenic peptide fragment SEQ ID NO:17.  
 XX  
 KW Human; CASB933; immunogenic; cytostatic; immunosuppressive; gene therapy;  
 KW vaccine; cytostatic; immunosuppressive; gene therapy; vaccine; medicine;  
 KW preneoplastic lesion; lung cancer; small cell lung carcinoma; carcinoma;  
 KW non-small cell lung carcinoma; squamous carcinoma; epidermoid;  
 KW adenocarcinoma; bronchoalveolar carcinoma; large cell carcinoma;  
 KW bronchial gland tumour; mesothelioma; autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02003016344-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 12-AUG-2002; 2002WO-EP009006.  
 XX  
 PR 14-AUG-2001; 2001GB-00019823.  
 XX  
 PA (GLAX) GLAXOSMITHKLINE BIOLOGICALS SA.  
 XX  
 PI Gaulis SRJ, Vinals Y De BassolsC;  
 XX  
 DR WPI; 2003-342415/32.  
 XX  
 PT An immunogenic composition comprising a CASB933 polynucleotide,  
 PT polypeptide or its immunogenic fragment, for preventing or treating a  
 PT patient suffering from or susceptible to preneoplastic lesions of lung  
 PT cancer or lung cancer.  
 XX  
 PS Claim 5; Page 66; 109pp; English.  
 XX  
 CC The present invention describes an immunogenic composition (I) comprising  
 CC a CASB933 polynucleotide, polypeptide or its immunogenic fragment, or an  
 CC antigen presenting cell modified to express the polypeptide, and a  
 CC pharmaceutical carrier. (I) has cytostatic and immunosuppressive  
 CC activities, and can be used in gene therapy and vaccines. The immunogenic  
 CC composition is useful in medicine. The CASB933 polypeptides or its  
 CC immunogenic fragments, or CASB933 polynucleotides encoding the  
 CC polypeptides are useful for manufacturing an immunogenic composition for  
 CC preventing or treating a patient suffering from or susceptible to  
 CC preneoplastic lesions of lung cancer or lung cancer, such as small cell  
 CC lung carcinoma, non-small cell lung carcinoma, e.g. squamous (epidermoid)  
 CC carcinoma, adenocarcinoma including bronchoalveolar or large cell  
 CC (undifferentiated) carcinoma, carcinoids, bronchial gland tumours, or  
 CC mesotheliomas. The CASB933 polynucleotides and polypeptides are also  
 CC useful for diagnosing the presence of, or a susceptibility to,  
 CC preneoplastic lesions of lung cancer or lung cancer in a subject, or for  
 CC treating autoimmune diseases. The present sequence represents an  
 CC immunogenic peptide fragment of human CASB933, which is given in the  
 CC exemplification of the present invention

XX SQ Sequence 9 AA;  
Query Match 1.8%; Score 6; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 64 EAVAQL 69  
Db 4 EAVAQL 9  
RESULT 166  
ABU03359  
ID ABU03359 standard; protein; 9 AA.  
XX AC ABU03359;  
XX DT 29-JAN-2003 (first entry)  
XX DE Human expressed protein tag (EPT) #139.  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX KW protease; protease inhibitor; transporter; cytoskeletal protein;  
XX KW receptor; transcription factor; cancer; MHC;  
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX OS Homo sapiens.  
XX DN WO200278524-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-US009671.  
XX PR 28-MAR-2001; 2001US-0279495P.  
XX PR 21-MAY-2001; 2001US-0292544P.  
XX PR 08-AUG-2001; 2001US-0310801P.  
XX PR 01-OCT-2001; 2001US-0326370P.  
XX PR 04-DEC-2001; 2001US-0336780P.  
XX PR 20-FEB-2002; 2002US-0358985P.  
XX PA (ZYCO-) ZYCOS INC.  
XX PI Chicx RM, Tomlinson AZ, Urban RG;  
XX WPI; 2003-040607/03.  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX Claim 10; SEQ ID NO 139; 134pp; English.  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_ppt\_sequences

XX SQ Sequence 9 AA;  
Query Match 1.8%; Score 6; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 230 KLIGEL 235  
Db 1 KLIGEL 6  
RESULT 167  
AAW71167  
ID AAW71167 standard; peptide; 10 AA.  
XX AC AAW71167;  
XX DT 26-OCT-1998 (first entry)  
XX DE Peptide used in an absorption material for hyperlipidemia therapy.  
XX KW Absorption material; hyperlipidemia therapy; removal; LDL; blood.  
XX OS Synthetic.  
XX PN JP10179733-A.  
XX PD 07-JUL-1998.  
XX PF 20-DEC-1996; 96JP-00354650.  
XX PR 20-DEC-1996; 96JP-00354650.  
XX PA (ASAH ) ASahi Kasei Kogyo KK.  
XX PA (ASAH ) ASahi Medical Co Ltd.  
XX WPI; 1998-440253/38.  
XX Absorption material for hyperlipidaemia therapy - comprises peptide  
PT containing amino acid satisfying predetermined formulae.  
XX Example 2; Page 9; 11pp; Japanese.  
XX AAW71164-67 represent peptides of electric charge used in an absorption  
CC material for hyperlipidemia therapy. The material enables the safe  
CC removal of LDL from blood  
XX Sequence 10 AA;  
Query Match 1.8%; Score 6; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 165 FRDFFK 170  
Db 3 FRDFFK 8  
RESULT 168  
AAG70315  
ID AAG70315 standard; peptide; 11 AA.  
XX AC AAG70315;  
XX DT 17-JUL-2001 (first entry)  
XX DE Human Chk2 kinase inhibitory peptide #1936.  
XX KW Human; Cdc25C; Chk1; Chk2; G2 checkpoint; cell cycle regulation; cancer;  
XX KW cell proliferation; apoptosis.  
XX OS Synthetic.

XX PN WO200121771-A2.  
 XX PD 29-MAR-2001.  
 XX PF 21-SEP-2000; 2000WO-IB001438.  
 XX PR 22-SEP-1999; 99JP-00269398.  
 XX PR 30-NOV-1999; 99JP-00340322.  
 XX PA (CANB-) CANBAS CO LTD.  
 XX PI Suganuma M, Kawabe T;  
 XX DR WPI; 2001-343125/36.  
 XX PT Isolated or recombinant polypeptide of 7-11 amino acids, useful for  
 PT treating cell proliferative disorders, e.g. to stop the growth of, or  
 PT kill cancer cells, by disrupting the G2 cell cycle arrest checkpoint.  
 XX PS Example 2; Page 97; 126pp; English.  
 XX CC The present invention describes a number of peptides which disrupt the G2  
 CC cell cycle checkpoint when administered to a cell. They act by inhibiting  
 CC Cdk1 and Cdk2 kinases, and may be derived from Cdc25C. The peptides, an  
 CC example of which is shown here, are useful in the treatment of cell  
 CC proliferation diseases, such as cancer, as the inhibition of the Chks  
 CC allows DNA damage and induces apoptosis  
 XX SQ Sequence 11 AA;  
 XX Query Match 1.8%; Score 6; DB 4; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 2e+02;  
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 250 SKPENL 255  
 DB |||||  
 6 SKPENL 11  
 RESULT 169  
 ABG74903  
 ID ABG74903 standard; peptide; 11 AA.  
 XX AC ABG74903;  
 XX DT 04-JUL-2003 (first entry)  
 XX DE Human TFR2-alpha peptide fragment #2.  
 XX KW Human; TFR2-alpha; diagnosis; haemochromatosis; transferrin receptor-2.  
 XX OS Homo sapiens.  
 XX PN WO2003016557-A1.  
 XX PD 27-FEB-2003.  
 XX PF 19-AUG-2002; 2002WO-AT000250.  
 XX PR 17-AUG-2001; 2001AT-00001295.  
 XX PA (VIEN-) VIENNALAB LABORDIAGNOSTIKA GMBH.  
 XX PI Camaschella C, Kury F, Oberkanins C;  
 XX DR WPI; 2003-268341/26.  
 XX PT Diagnosis of hemochromatosis, by detecting a specific deletion in the  
 PT transferrin receptor-2 gene, or corresponding deletion from the encoded  
 PT protein.  
 XX PS Claim 3; Page 3; 23pp; German.

XX CC This invention describes a novel method for diagnosing haemochromatosis  
 CC by testing a biological sample for a sequence that represents a 12  
 CC nucleotide deletion from exon 16 of the cDNA for TFR2 (transferrin  
 CC receptor-2). The method can detect haemochromatosis before clinical  
 CC symptoms are manifest (by which time irreversible damage to organs has  
 CC occurred). This sequence represents a fragment of the human TFR2-alpha  
 CC (see Genbank NM 003227) which is used in the method described in the  
 XX disclosure of the invention  
 XX SQ Sequence 11 AA;  
 XX Query Match 1.8%; Score 6; DB 6; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 2e+02;  
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 65 AVAQLA 70  
 DB |||||  
 2 AVAQLA 7  
 RESULT 170  
 AAU69193  
 ID AAU69193 standard; peptide; 14 AA.  
 XX AC AAU69193;  
 XX DT 29-JAN-2002 (first entry)  
 XX DE Human Acetylcholine receptor alpha subunit peptide AChR 330-343.  
 XX KW Human; epitope; autoimmune disease; myasthenia gravis;  
 KW Human leukocyte antigen; acetylcholine receptor; HLA DR3; HLA DR2; AChR;  
 KW antigen; immunosuppressive; major histocompatibility complex; MHC.  
 XX OS Homo sapiens.  
 XX PN WO200174848-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US010450.  
 XX PR 31-MAR-2000; 2000US-0193745P.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Deshpande S, Spack E, Wehner N, Arimilli S;  
 XX DR WPI; 2001-648547/74.  
 XX PT Peptide epitopes of the acetylcholine receptor target helper T cells  
 PT recognize an antigen in association with an MHC component and are useful  
 PT to treat autoimmune disease particularly myasthenia gravis.  
 XX PS Example 1; Fig 2; 46pp; English.  
 XX CC The invention relates to a composition comprising an isolated  
 CC acetylcholine receptor (AChR) oligopeptide of about 12 to 20 amino acids.  
 CC The peptides form a set of 69 overlapping antigenic peptide epitopes  
 CC which show various affinities for human leukocyte antigens (HLA) HLA-DR2  
 CC and DR3. Also included is a composition comprising an antigenic peptide  
 CC and a Major histocompatibility complex (MHC) component having an  
 CC antigenic binding site, where binding of the peptide to the binding site  
 CC induces non-responsiveness in a target T cell in a mammal, where the MHC  
 CC component is an MHC class II component. Peptides with affinity for HLA-  
 CC DR2 and DR-3 from proteins other than AChR are also included. The  
 CC composition is used to treat myasthenia gravis and other autoimmune  
 CC diseases. The present sequence is an antigenic peptide epitope  
 XX SQ Sequence 14 AA;  
 XX Query Match 1.8%; Score 6; DB 4; Length 14;





PD 02-MAR-1999.  
XX  
PF 02-MAY-1995; 95US-00432871.  
XX  
PR 02-MAY-1994; 94US-00237592.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Black ME, Loeb LA;  
XX  
DR WPI; 1999-189650/16.  
XX  
PT New Herpesviridae thymidine kinase mutant nucleic acids - used to  
develop products for treating e.g. tumours, autoimmune diseases,  
PT allergies, restenosis or viral, bacterial or parasitic diseases.  
XX  
PS Disclosure; Fig 15; 72pp; English.  
XX  
CC This sequence represents a mutant form of amino acids 159-173 of the  
herpes simplex virus type 1 (HSV-1) thymidine kinase (TK) protein which  
was used to generate mutant TK proteins. The invention relates to the  
CC generation of novel HSV-1 TK or guanylate kinase (GK) genes with a  
mutation upstream, within or downstream from a DRH nucleoside binding  
CC site. The TK enzymes can be used for inhibiting pathogenic agents, e.g.  
tumours, hyperkeratosis, psoriasis, prostate hypertrophy,  
CC hyperthyroidism, endocrinopathies, autoimmune diseases, allergies,  
CC restenosis, viral diseases such as AIDS, hepatitis, intracellular  
CC parasitic diseases or bacterial infection  
XX  
SQ Sequence 15 AA;  
Query Match 1.8%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 235 LILDRH 240  
Db 1 LILDRH 6  
RESULT 174  
AAW971150  
ID AAW971150 standard; peptide; 15 AA.  
XX  
AC AAW971150;  
XX  
DT 22-JUN-1999 (first entry)  
XX  
DE HSV-1 TK amino acids 159-173 mutant #11.  
XX  
KW HSV-1; thymidine kinase; mutation; DRH nucleoside binding site; enzyme;  
pathogen; tumour; hyperkeratosis; psoriasis; prostate hypertrophy;  
KW hyperthyroidism; endocrinopathy; autoimmune disease; allergy; restenosis;  
KW viral disease; AIDS; hepatitis; parasitic; bacterial infection.  
XX  
OS Herpes simplex virus unknown type.  
OS Synthetic.  
XX  
PN US5877010-A.  
XX  
PD 02-MAR-1999.  
XX  
PF 02-MAY-1995; 95US-00432871.  
XX  
PR 02-MAY-1994; 94US-00237592.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Black ME, Loeb LA;  
XX  
DR WPI; 1999-189650/16.  
XX  
PT New Herpesviridae thymidine kinase mutant nucleic acids - used to

PT develop products for treating e.g. tumours, autoimmune diseases,  
PT allergies, restenosis or viral, bacterial or parasitic diseases.  
XX  
PS Disclosure; Fig 15; 72pp; English.  
XX  
CC This sequence represents a mutant form of amino acids 159-173 of the  
herpes simplex virus type 1 (HSV-1) thymidine kinase (TK) protein which  
was used to generate mutant TK proteins. The invention relates to the  
CC generation of novel HSV-1 TK or guanylate kinase (GK) genes with a  
mutation upstream, within or downstream from a DRH nucleoside binding  
CC site. The TK enzymes can be used for inhibiting pathogenic agents, e.g.  
tumours, hyperkeratosis, psoriasis, prostate hypertrophy,  
CC hyperthyroidism, endocrinopathies, autoimmune diseases, allergies,  
CC restenosis, viral diseases such as AIDS, hepatitis, intracellular  
CC parasitic diseases or bacterial infection  
XX  
SQ Sequence 15 AA;  
Query Match 1.8%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 235 LILDRH 240  
Db 1 LILDRH 6  
RESULT 175  
ABG74902  
ID ABG74902 standard; peptide; 15 AA.  
XX  
AC ABG74902;  
XX  
DT 04-JUL-2003 (first entry)  
XX  
DE Human TFR2-alpha peptide fragment #1.  
XX  
KW Human; TFR2-alpha; diagnosis; haemochromatosis; transferrin receptor-2.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016557-A1.  
XX  
PD 27-FEB-2003.  
XX  
PF 19-AUG-2002; 2002WO-AF000250.  
XX  
PR 17-AUG-2001; 2001AT-00001295.  
XX  
PA (VIEN-) VIENNALAB LABORDIAGNOSTIKA GMBH.  
XX  
PI Camaschella C, Kury F, Oberkanins C;  
XX  
DR WPI; 2003-368341/26.  
XX  
PT Diagnosis of hemochromatosis, by detecting a specific deletion in the  
transferrin receptor-2 gene, or corresponding deletion from the encoded  
PT protein.  
XX  
PS Claim 3; Page 3; 23pp; German.  
XX  
CC This invention describes a novel method for diagnosing haemochromatosis  
by testing a biological sample for a sequence that represents a 12  
CC nucleotide deletion from exon 16 of the cDNA for TFR2 (transferrin  
receptor-2). The method can detect haemochromatosis before clinical  
CC symptoms are manifest (by which time irreversible damage to organs has  
occurred). This sequence represents a fragment of the human TFR2-alpha  
CC (see Genbank NM 003227) which is used in the method described in the  
disclosure of the invention  
XX  
SQ Sequence 15 AA;  
Query Match 1.8%; Score 6; DB 6; Length 15;



CC Gene within a cell, or to replace a specific gene which is defective in  
 CC proper expression using gene therapy, e.g. including adenosine deaminase  
 CC deficiency, and Alzheimer's diseases. The mutants are utilised as a  
 CC conditionally lethal marker for homologous recombination. This is the  
 CC amino acid sequence of a herpesviridae thymidine kinase (TK) mutant  
 CC peptide  
 XX  
 SQ Sequence 15 AA;

Query Match 1.8%; Score 6; DB 6; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 LILDRH 240  
 Db 1 LILDRH 6

RESULT 178  
 AAM22114  
 ID AAM22114 standard; protein; 18 AA.  
 XX  
 AC AAM22114;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #8548 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.

30-JAN-2001; 2001WO-US000670.  
 04-FEB-2000; 2000US-0180312P.  
 26-MAY-2000; 2000US-0207456P.  
 30-JUN-2000; 2000US-00608408.  
 03-AUG-2000; 2000US-00632366.  
 21-SEP-2000; 2000US-0234687P.  
 27-SEP-2000; 2000US-0236359P.  
 04-OCT-2000; 2000GB-00024263.  
 (MOLE-) MOLECULAR DYNAMICS INC.  
 Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-488901/53.  
 Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human cervical epithelial cells.

Claim 27; SEQ ID NO 26940; 487pp; English.  
 The present invention relates to human single exon nucleic acid probes  
 (SENPs: see AAI10069-AAI28459). The present sequence is a peptide encoded  
 by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 can be used to produce a single exon microarray, which can be used for  
 measuring human gene expression in a sample derived from human cervical  
 epithelial cells. By measuring gene expression, the probes are therefore  
 useful in grading and/or staging of diseases of the cervix, notably  
 cervical cancer. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 18 AA;

Query Match 1.8%; Score 6; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 250 SKPENL 255  
 Db 11 SKPENL 16

RESULT 179  
 ABB44520  
 ID ABB44520 standard; peptide; 18 AA.  
 XX  
 AC ABB44520;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 DE Peptide #12026 encoded by human foetal liver single exon probe.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.

30-JAN-2001; 2001WO-US000669.  
 04-FEB-2000; 2000US-0180312P.  
 26-MAY-2000; 2000US-0207456P.  
 30-JUN-2000; 2000US-00608408.  
 03-AUG-2000; 2000US-00632366.  
 21-SEP-2000; 2000US-0234687P.  
 27-SEP-2000; 2000US-0236359P.  
 04-OCT-2000; 2000GB-00024263.  
 (MOLE-) MOLECULAR DYNAMICS INC.  
 Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-483447/52.  
 Human genome-derived single exon nucleic acid probes useful for analyzing  
 gene expression in human fetal liver.

Claim 27; SEQ ID NO 37155; 639pp + Sequence Listing; English.  
 The invention relates to a single exon nucleic acid probe for measuring  
 human gene expression in a sample derived from human foetal liver. The  
 single exon nucleic acid probes may be used for predicting, measuring and  
 displaying gene expression in samples derived from human fetal liver. The  
 present sequence is a peptide encoded by a single exon nucleic acid probe  
 of the invention. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Sequence 18 AA;

Query Match 1.8%; Score 6; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SKPENL 255  
 Db 11 SKPENL 16

RESULT 180  
 AAM38567  
 ID AAM38567 standard; protein; 18 AA.  
 XX  
 AC AAM38567;  
 XX  
 DT 17-OCT-2001 (first entry)

XX DE Peptide #12604 encoded by probe for measuring placental gene expression.  
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
 XX KW genetic disorder.  
 OS Homo sapiens.  
 XX OS  
 PN WO200157272-A2.  
 XX PD  
 XX PF 09-AUG-2001.  
 XX PF  
 XX PF 30-JAN-2001; 2001WO-US000663.  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-488897/53.  
 XX DR  
 XX DR Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX PT gene expression in human placenta.  
 XX XX  
 XX PS Claim 27; SEQ ID NO 38836; 654pp; English.  
 XX CC The present invention relates to single exon nucleic acid probes (SENP;  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 XX CC  
 XX SQ Sequence 18 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 250 SKPENL 255  
 Db 11 SKPENL 16  
 |||||  
 RESULT 181  
 AAM78321  
 ID AAM78321 standard; protein; 18 AA.  
 XX AC  
 XX AC AAM78321;  
 XX XX  
 XX DT 06-NOV-2001 (first entry)  
 XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38627.  
 XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 XX KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX OS Homo sapiens.  
 XX OS  
 XX PN WO200157276-A2.  
 XX XX  
 XX PD 09-AUG-2001.  
 XX PF  
 XX PF 30-JAN-2001; 2001WO-US000668.  
 XX XX  
 XX PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX XX  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-488900/53.  
 XX DR  
 XX DR Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX PT gene expression in human bone marrow.  
 XX XX  
 XX PS Example 4; SEQ ID NO 38627; 658pp + Sequence Listing; English.  
 XX CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention  
 XX CC  
 XX SQ Sequence 18 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 250 SKPENL 255  
 Db 11 SKPENL 16  
 |||||  
 RESULT 182  
 AAM65705  
 ID AAM65705 standard; protein; 18 AA.  
 XX AC  
 XX AC AAM65705;  
 XX XX  
 XX DT 05-NOV-2001 (first entry)  
 XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37810.  
 XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
 XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
 XX OS Homo sapiens.  
 XX OS  
 XX PN WO200157275-A2.  
 XX XX  
 XX PD 09-AUG-2001.  
 XX PF  
 XX PF 30-JAN-2001; 2001WO-US000667.  
 XX XX  
 XX PR 04-FEB-2000; 2000US-0180312P.  
 XX PR 26-MAY-2000; 2000US-0207456P.  
 XX PR 30-JUN-2000; 2000US-00608408.  
 XX PR 03-AUG-2000; 2000US-00632366.  
 XX PR 21-SEP-2000; 2000US-0234687P.  
 XX PR 27-SEP-2000; 2000US-0236359P.  
 XX PR 04-OCT-2000; 2000GB-00024263.  
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-483446/52.  
 XX DR  
 XX XX  
 XX PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.



CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagenet syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a peptide/protein encoded by a single exon probe of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 CC SQ Sequence 18 AA;

Query Match 1.8%; Score 6; DB 5; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 250 SKPENL 255  
 DB 11 SKPENL 16  
 |||||

RESULT 185  
 AAR54884  
 ID AAR54884 standard; peptide; 22 AA.  
 AC AAR54884;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 03-NOV-1994 (first entry)  
 XX  
 XX Ion channel forming amphiphilic peptide.  
 XX  
 XX Ionophore; antimicrobial; antiviral; antibacterial; antiparasitic;  
 XX spermicide; wound healing; burns; anticancer; preservative; sterilant;  
 XX disinfectant; plant protection.  
 XX  
 XX Synthetic.  
 XX  
 XX WO9409810-A1.  
 XX  
 XX 11-MAY-1994.  
 XX  
 XX 22-OCT-1993; 93WO-US010337.  
 XX  
 XX 26-OCT-1992; 92US-00965663.  
 XX  
 XX (MAGA-) MAGAININ PHARM INC.  
 XX  
 XX Maloy WL; Kari UP;  
 XX  
 XX WPI; 1994-167120/20.  
 XX  
 XX New ion channel forming amphiphilic - useful as antimicrobial,  
 XX antitumour, antiparasitic and spermicidal agents.  
 XX

PS Claim 6; Page 37; 43pp; English.  
 XX  
 CC The peptide sequence is that of an ion forming peptide used to inhibit  
 CC the growth of target cells, viruses and virally infected cells in a host,  
 CC i.e. they are antimicrobial, antiviral, antibacterial, anticancer and  
 CC antiparasitic agents or spermicides. They can also be used to stimulate  
 CC wound healing and can be used to treat burns. The peptides can be used in  
 CC human or veterinary medicine as preservatives, sterilants or  
 CC disinfectants and in plant protection. See also AAR54880-906. (Updated on  
 CC 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 22 AA;  
 Query Match 1.8%; Score 6; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 IVKILK 22  
 DB 16 IVKILK 21  
 |||||  
 RESULT 186  
 AAW66527  
 ID AAW66527 standard; peptide; 22 AA.  
 AC AAW66527;  
 XX  
 DT 25-NOV-1998 (first entry)  
 DT  
 XX Amphiphilic peptide.  
 DE  
 XX magainin; analogue; antimicrobial; antitumour; wound healing; CPF;  
 KW amphiphilic; XEP peptide.  
 KW  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1..22 /note= "All residues except Gly are D-form residues"  
 FT  
 XX US5792831-A.  
 XX  
 XX 11-AUG-1998.  
 XX  
 XX 17-NOV-1994; 94US-00343882.  
 XX  
 XX 08-FEB-1990; 90US-00476629.  
 XX 14-MAY-1990; 90US-00522688.  
 XX 28-APR-1992; 92US-00874685.  
 XX 05-OCT-1993; 93US-00133740.  
 XX  
 XX (MAGA-) MAGAININ PHARM INC.  
 XX  
 XX Maloy WL;  
 XX  
 XX WPI; 1998-456190/39.  
 XX  
 XX Magainin peptide analogues - useful as antimicrobial or antitumour  
 XX agents, etc.  
 XX  
 XX Disclosure; Col 21; 25pp; English.  
 XX  
 CC The invention relates to analogues of a magainin I or II, D-form  
 CC analogues, deletion analogues or related peptides. It also relates to  
 CC basic polypeptides having at least 16 amino acids, including at least 8  
 CC hydrophobic amino acids and at least 8 hydrophilic amino acids. The  
 CC peptides may be used as antimicrobial agents, antiviral agents,  
 CC antibiotics, antitumour agents, antiparasitic agents, spermicides,  
 CC preservatives or sterilants, or agents for promoting wound healing. The  
 CC present sequence represents a specific example of a peptide disclosed in  
 CC the specification  
 CC

SQ Sequence 22 AA;

Query Match 1.8%; Score 6; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILK 22  
 |||||  
 Db 16 IVKILK 21

RESULT 187

ABO12236  
 ID ABO12236 standard; peptide; 23 AA.

XX

AC ABO12236;

XX 25-AUG-2003 (first entry)

DT

XX Human zinc finger DNA binding domain #535.

XX

KW Composite binding polypeptide; zinc finger nucleic acid binding domain;  
 XW autoimmune disorder; immunosuppressive; zinc finger DNA binding domain;  
 KW human.

XX

OS Homo sapiens.

XX

PN W0200299084-A2.

XX

PD 12-DEC-2002.

XX

XX 04-APR-2002; 2002WO-US022272.

XX

XX 04-APR-2001; 2001GB-00008491.

XX

PA (SANG-) SANGAMO BIOSCIENCES INC.

XX

PI Moore M, Sepp A, Isalan M, Choo Y;

XX

DR WPI; 2003-278214/27.

XX

PT New composite binding zinc finger polypeptide, useful for designing  
 PT sequence-specific binding proteins regulating gene expression in the  
 PT fields of molecular biology, and for the diagnosis and treatment of  
 PT autoimmune disorders.

XX

PS Example 2; Page 82; 157pp; English.

XX

CC The invention relates to a composite binding polypeptide comprising a  
 CC first natural binding domain derived from a first natural binding  
 CC polypeptide and a second natural binding domain derived from a second  
 CC natural binding polypeptide, where the first and second natural binding  
 CC polypeptides may be the same or different and where the polypeptide binds  
 CC to a target differing from the natural target of both the first and  
 CC second binding polypeptides. The invention also relates to a chimeric  
 CC polypeptide comprising a binding polypeptide cited above and a biological  
 CC effector domain, a library of natural binding domains, a library of  
 CC natural zinc finger nucleic acid binding domains comprising a linker  
 CC attached to it, a method for selecting a binding polypeptide capable of  
 CC binding to a target site and a method for designing a composite binding  
 CC polypeptide. The methods and compositions of the present invention are  
 CC useful for designing sequence-specific binding proteins for regulation of  
 CC gene expression in the fields of molecular biology. They can also be used  
 CC for the diagnosis and treatment of autoimmune disorders, and as research  
 CC tools and in transgenic animals. This sequence represents a human zinc  
 CC finger DNA binding domain used in the scope of the invention

SQ Sequence 23 AA;

Query Match

Best Local Similarity 1.8%; Score 6; DB 6; Length 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 LTRHKV 196  
 |||||  
 Db 16 LTRHKV 21

RESULT 188

ABO11924

ID ABO11924 standard; peptide; 23 AA.

XX

AC ABO11924;

XX

DT 25-AUG-2003 (first entry)

XX

DE Human zinc finger DNA binding domain #224.

XX

KW Composite binding polypeptide; zinc finger nucleic acid binding domain;  
 XW autoimmune disorder; immunosuppressive; zinc finger DNA binding domain;  
 KW human.

XX

OS Homo sapiens.

XX

PN W0200299084-A2.

XX

PD 12-DEC-2002.

XX

XX 04-APR-2002; 2002WO-US022272.

XX

XX 04-APR-2001; 2001GB-00008491.

XX

PA (SANG-) SANGAMO BIOSCIENCES INC.

XX

PI Moore M, Sepp A, Isalan M, Choo Y;

XX

DR WPI; 2003-278214/27.

XX

PT New composite binding zinc finger polypeptide, useful for designing  
 PT sequence-specific binding proteins regulating gene expression in the  
 PT fields of molecular biology, and for the diagnosis and treatment of  
 PT autoimmune disorders.

XX

PS Example 1; Page 75; 157pp; English.

XX

CC The invention relates to a composite binding polypeptide comprising a  
 CC first natural binding domain derived from a first natural binding  
 CC polypeptide and a second natural binding domain derived from a second  
 CC natural binding polypeptide, where the first and second natural binding  
 CC polypeptides may be the same or different and where the polypeptide binds  
 CC to a target differing from the natural target of both the first and  
 CC second binding polypeptides. The invention also relates to a chimeric  
 CC polypeptide comprising a binding polypeptide cited above and a biological  
 CC effector domain, a library of natural binding domains, a library of  
 CC natural zinc finger nucleic acid binding domains comprising a linker  
 CC attached to it, a method for selecting a binding polypeptide capable of  
 CC binding to a target site and a method for designing a composite binding  
 CC polypeptide. The methods and compositions of the present invention are  
 CC useful for designing sequence-specific binding proteins for regulation of  
 CC gene expression in the fields of molecular biology. They can also be used  
 CC for the diagnosis and treatment of autoimmune disorders, and as research  
 CC tools and in transgenic animals. This sequence represents a human zinc  
 CC finger DNA binding domain used in the scope of the invention

SQ Sequence 23 AA;

Query Match

Best Local Similarity 1.8%; Score 6; DB 6; Length 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 LTRHKV 196  
 |||||  
 Db 16 LTRHKV 21

RESULT 189



AD47956  
ID ADB47956 standard; protein; 24 AA.  
XX  
AC ADB47956;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Novel human secreted protein associated polypeptide #35.  
XX  
KW human; secreted protein; insulin; haemoglobin S; haemoglobin B;  
KW superoxide; SOD; catalase; DNA repair protein; oncogene;  
KW tumour suppressor; tumour necrosis factor; TNF; inflammation;  
KW blood vessel growth inhibition; immune response; immune system disorder;  
KW hyperproliferative disease; neoplasm; cardiovascular disorder;  
KW peripheral artery disease; limb ischaemia; cardio-arterial fistula;  
KW arteriovenous fistula; congenital heart defect;  
KW neovascularisation disorder; wound healing;  
KW epithelial cell proliferation; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; mania; dementia;  
KW infectious disease.  
XX  
OS Homo sapiens.  
XX  
XX US2003054443-A1.  
XX  
PD 20-MAR-2003.  
XX  
XX 04-OCT-2001; 2001US-00969730.  
XX  
PR 05-AUG-1997; 97US-0054798P.  
PR 05-AUG-1997; 97US-0054803P.  
PR 05-AUG-1997; 97US-0054804P.  
PR 05-AUG-1997; 97US-0054806P.  
PR 05-AUG-1997; 97US-0054807P.  
PR 05-AUG-1997; 97US-0054808P.  
PR 05-AUG-1997; 97US-0054809P.  
PR 05-AUG-1997; 97US-0055309P.  
PR 05-AUG-1997; 97US-0055310P.  
PR 05-AUG-1997; 97US-0055311P.  
PR 05-AUG-1997; 97US-0055312P.  
PR 05-AUG-1997; 97US-0055366P.  
PR 18-AUG-1997; 97US-0055970P.  
PR 18-AUG-1997; 97US-0055986P.  
PR 19-AUG-1997; 97US-0056364P.  
PR 19-AUG-1997; 97US-0056365P.  
PR 19-AUG-1997; 97US-0056366P.  
PR 19-AUG-1997; 97US-0056387P.  
PR 19-AUG-1997; 97US-0056370P.  
PR 19-AUG-1997; 97US-0056371P.  
PR 19-AUG-1997; 97US-0056557P.  
PR 19-AUG-1997; 97US-0056563P.  
PR 19-AUG-1997; 97US-0056731P.  
PR 19-AUG-1997; 97US-0056732P.  
PR 04-AUG-1998; 98WO-US016235.  
PR 04-FEB-1999; 99US-00244112.  
PR 06-OCT-2000; 2000US-0238291P.  
PR 01-FEB-2001; 2001US-00774639.  
XX  
PA (RUBE/) RUBEN S. M.  
PA (SOPP/) SOPPET D. R.  
PA (EBNE/) EBNER R.  
PA (OLSE/) OLSEN H. S.  
PA (YOUN/) YOUNG P. E.  
PA (GREE/) GREENE J. M.  
PA (FERE/) FERRIE A. M.  
PA (YUGG/) YU G.  
PA (NIJ/) NI J.  
PA (ROSE/) ROSEN C. A.  
PA (BREW/) BREWER L. A.  
PA (JANA/) JANAT F.  
PA (BIRS/) BIRSE C. E.  
XX  
XX Ruben SM, Soppet DR, Ebner R, Olsen HS, Young PE, Greene JM;  
PI

PI Ferrie AM, Yu G, Ni J, Rosen CA, Brewer LA, Janat F, Birse CE;  
XX WPI; 2003-695903/66.  
XX  
PT Novel human secreted proteins useful for treating and/or diagnosing  
PT disorders of immune system, cardiovascular disorders such as peripheral  
PT artery disease, neurological diseases such as Alzheimer's disease.  
XX  
XX Disclosure; Page 15; 33pp; English.  
XX  
CC The invention relates to novel human secreted proteins. The protein is  
CC useful for preventing, treating or ameliorating a medical condition. The  
CC protein is useful for diagnosing a pathological condition or  
CC susceptibility to a pathological condition in a subject. The protein is  
CC useful for identifying a binding partner. The nucleic acid is useful for  
CC diagnosing pathological condition or a susceptibility to pathological  
CC condition in a subject. The protein is useful as reagents for  
CC differential identification of the tissues or cell types present in a  
CC biological sample. The protein can be administered to patients having  
CC absent or decreased levels of polypeptides e.g. insulin, to supplement  
CC absent or decreased levels of different polypeptides, e.g. haemoglobin S  
CC for haemoglobin B, superoxide (SOD), catalase, DNA repair protein, to  
CC inhibit the activity of a polypeptide e.g. an oncogene or tumour  
CC suppressor, to activate the activity of membrane bound receptor by competing  
CC with it for free ligand e.g. soluble tumour necrosis factor (TNF)  
CC receptors used in reducing inflammation, or to bring about a desired  
CC response e.g. blood vessel growth inhibition, enhancement of immune  
CC response to proliferative cells or tissues. The protein and the nucleic  
CC acid are useful for treating, preventing, detecting, diagnosing disorders  
CC of immune system involving abnormal growth of specific types of cells as  
CC well as of other cell types where expression has been observed. The  
CC protein, the nucleic acid and antibodies are useful for treating,  
CC preventing and/or diagnosing diseases, disorders and/or conditions of  
CC immune system, hyperproliferative disorders including neoplasms,  
CC cardiovascular disorders (such as peripheral artery disease, limb  
CC ischaemia, arterio-arterial fistula, arteriovenous fistula, congenital  
CC heart defects, etc), neovascularisation disorders, wound healing and  
CC epithelial cell proliferation, neurological diseases (such as Alzheimer's  
CC disease, Parkinson's disease, Huntington's disease, mania, dementia,  
CC etc), infectious diseases caused by virus, bacteria, fungi, etc. The  
CC present sequence represents the amino acid sequence of a novel human  
CC secreted protein associated polypeptide.  
XX  
SQ Sequence 24 AA;  
XX  
Query Match 1.8%; Score 6; DB 7; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 33 KKTDKA 38  
DB 8 KKTDKA 13  
RESULT 190  
ABBI6060  
ID ABB16060 standard; protein; 28 AA.  
XX  
AC ABB16060;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polypeptide SEQ ID NO 4717.  
XX  
KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX

OS Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180638P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-01981123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218280P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226688P.

XX 22-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0242221P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249278P.

PR 17-NOV-2000; 2000US-0249299P.

PR 01-DEC-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.

PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA

XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-541565/60.  
 DR N-PSDB; ABA12386.  
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system cancers  
 PT and metastases.  
 XX Claim 11; SEQ ID NO 4717; 1701pp + Sequence Listing; English.  
 XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
 CC (ABA14678-AB18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
 XX SQ Sequence 28 AA;  
 Query Match 1.8%; Score 6; DB 4; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 271 FEAPHV 276  
 Db 20 FEAPHV 25  
 RESULT 191  
 ABG76643  
 ID ABG76643 standard; peptide; 29 AA.  
 AC ABG76643;  
 XX 05-NOV-2002 (first entry)  
 DT Human SOUP1 maximal transmembrane domain peptide MB-5.  
 DE Suppressor of uncoupling protein 1; SOUP1; membrane stability; obesity;  
 KW organellar function; mitochondria; peroxisome; cellular metabolism; ROS;  
 KW homeostasis; body weight; metabolic disorder; adipositas; cachexia;  
 KW eating disorder; wasting syndrome; mitochondrial disorder; antilipaeamic;  
 KW pancreatic dysfunction; anorectic; immunomodulator; gene therapy;  
 KW regulator of transporter molecule; modifier of mitochondrial protein;  
 KW fruit fly; human; mouse; transmembrane domain.  
 XX Homo sapiens.  
 OS WO200242455-A2.  
 XX 30-MAY-2002.  
 PD 23-NOV-2001; 2001WO-EP013663.  
 PF 23-NOV-2000; 2000EP-00125693.  
 PR (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.  
 XX Steuernagel A, Broenner G, Dohrmann C, Ciossek T, Wehr R;  
 PI Rudolph B, Rudolph D;

XX WPI; 2002-599380/64.  
 DR New suppressor of uncoupling protein 1 polynucleotide encoding a  
 XX polypeptide contributing to membrane stability and/or function of  
 PT organelles, useful for treating metabolic disorders and mitochondrial  
 PT disorders.  
 XX Claim 15; Fig 5e; 141pp; English.  
 XX The present invention relates to a new nucleic acid molecule, termed  
 CC suppressor of uncoupling protein 1 (SOUP1), encoding a polypeptide  
 CC contributing to membrane stability and/or function of organelles. The  
 CC invention is useful for monitoring and/or controlling the function of a  
 CC gene and/or gene product which is influenced and/or modified by the SOUP1  
 CC polypeptide, where the gene and/or gene product is expressed in  
 CC organelles such as mitochondria or peroxisome, or for identifying  
 CC substances capable of interacting with the polypeptide. The invention is  
 CC also useful for identifying a polypeptide or substance(s) involved in  
 CC cellular metabolism in an animal or capable of modifying homeostasis, or  
 CC for identifying a polypeptide involved in the regulation of body weight  
 CC in a mammal. The invention is further useful for detecting and/or  
 CC verifying a disorder in cells, cell masses, organs and/or subjects, and  
 CC for treating, alleviating and/or preventing a disorder in cells, cell  
 CC masses, organs and/or subjects, where the disorder is a metabolic  
 CC disorder such as obesity, adipositas, eating disorders, wasting syndromes  
 CC (cachexia), mitochondrial disorders, pancreatic dysfunctions, and  
 CC disorders related to ROS production. The invention is useful for the  
 CC detection of polypeptides capable of contributing to membrane stability  
 CC and/or function in organelles, capable of modifying mitochondrial  
 CC proteins and/or capable of inducing cellular metabolism. The present  
 CC amino acid sequence represents a SOUP1 transmembrane domain peptide of  
 CC the invention  
 XX SQ Sequence 29 AA;  
 Query Match 1.8%; Score 6; DB 5; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 115 TVEYIS 120  
 Db 4 TVEYIS 9  
 RESULT 192  
 AAW15122  
 ID AAW15122 standard; peptide; 33 AA.  
 XX AAW15122;  
 AC 04-MAR-1998 (first entry)  
 DT Corticotropin-releasing factor (CRF) analogue IV (3ala Astressin).  
 DE Corticotropin-releasing factor; CRF; analogue; agonist; antagonist;  
 XX lipophilic; RT-PCR; ligand binding activity; treatment; astressin;  
 KW depression; psychosomatic disease; therapeutic agent; cyclic.  
 KW Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1 /note= "D-form residue"  
 FT Modified-site 10 /label= Nle  
 FT Modified-site 19 /note= "Norleucine"  
 FT /note= "The side chain carboxyl group of Glu at position  
 FT 19 forms a lactam bridge with the side chain amino group  
 FT of Lys at position 22"  
 FT Modified-site 22 /note= "The side chain amino group of Lys at position 22"

FT forms a lactam bridge with the side chain amino group of  
 FT Glu at position 19"  
 FT 30  
 FT /label= Nle  
 FT /note= "Norleucine"  
 XX  
 XX WO9718238-A2.  
 XX  
 XX  
 PD 22-MAY-1997.  
 XX  
 PF 14-NOV-1996; 96WO-EP005010.  
 XX  
 PR 14-NOV-1995; 95EP-00117940.  
 XX  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 XX  
 PI Spiess J, Ruehmann A;  
 XX WPI; 1997-289228/26.  
 XX  
 XX Lipophilic corticotropin-releasing factor analogues - have increased  
 PT ability to pass through the blood-brain barrier and are useful as  
 PT therapeutic agents for treating e.g. stress and anxiety.  
 XX  
 PS Disclosure; Page 7; 14pp; English.  
 XX  
 CC This 3Ala Astressin is a lipophilic corticotropin-releasing factor (CRF)  
 CC analogue where some amino acids in the native h/CRF (Astressin) have  
 CC been replaced by alanine at positions 20, 25 and 32. The CRF analogue  
 CC retains the CRF ligand binding activity. CRF is thought to integrate  
 CC endocrine, autonomic, immune and behavioural responses to stress in the  
 CC Central nervous system (CNS) through regulation of hypothalamus-pituitary  
 CC -adrenal axis leading to glucocorticoids release after stress exposure,  
 CC in which signal transduction is mediated through receptors. Analogues of  
 CC CRF can imitate or block biological function of CRF (agonists or  
 CC antagonists respectively), making them useful in investigation of  
 CC peptidergic systems and as therapeutic agents e.g. in prevention or  
 CC treatment of stress, depression, anxiety and other psychosomatic  
 CC diseases. They can be used alone or with other agents and treatments. The  
 CC analogues can easily pass through the blood-brain barrier due to their  
 CC high affinity to CRF receptor and raised lipophilicity  
 XX  
 SQ Sequence 33 AA;  
 Query Match 1.8%; Score 6; DB 2; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 67 AQLAQE 72  
 Db 14 AQLAQE 19  
 |||||  
 |||||  
 RESULT 193  
 ID ABB43220  
 XX ABB43220 standard; peptide; 34 AA.  
 XX  
 AC ABB43220;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 XX Peptide #10726 encoded by human foetal liver single exon probe.  
 DE Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX Homo sapiens.  
 OS  
 XX WO200157277-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000669.  
 PF  
 XX WPI; 2001-48897/53.  
 XX

PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483447/52.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human fetal liver.  
 XX  
 PS Claim 27; SEQ ID NO 35855; 639pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 34 AA;  
 Query Match 1.8%; Score 6; DB 4; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 287 TQPIVE 232  
 Db 1 TQPIVE 6  
 |||||  
 |||||  
 RESULT 194  
 ID AAM37060  
 XX AAM37060 standard; protein; 34 AA.  
 XX  
 AC AAM37060;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #11097 encoded by probe for measuring placental gene expression.  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 XX genetic disorder.  
 XX Homo sapiens.  
 OS  
 XX WO200157272-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000663.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-48897/53.  
 XX

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.

XX Claim 27; SEQ ID NO 37329; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders

SQ Sequence 34 AA;

Query Match 1.8%; Score 6; DB 4; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TQPIVE 292  
 Db |||||  
 1 TQPIVE 6

RESULT 195  
 AAM76952  
 ID AAM76952 standard; protein; 34 AA.

AC AAM76952;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 37258.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.

XX Example 4; SEQ ID NO 37258; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention

XX Sequence 34 AA;

Query Match 1.8%; Score 6; DB 4; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TQPIVE 292  
 Db |||||  
 1 TQPIVE 6

RESULT 196  
 AAM64127  
 ID AAM64127 standard; protein; 34 AA.

XX AAM64127;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36232.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.

XX Example 4; SEQ ID NO 36232; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention

XX Sequence 34 AA;

Query Match 1.8%; Score 6; DB 4; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TQPIVE 292  
 Db |||||  
 1 TQPIVE 6

RESULT 197  
 ABG58614  
 ID ABG58614 standard; peptide; 34 AA.

XX ABG58614;

XX 25-FEB-2003 (first entry)  
DT Human liver peptide, SEQ ID No 37262.  
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
XX hypercholesterolaemia; coronary heart disease.  
KW Homo sapiens.  
OS  
XX WO200157273-A2.  
FN  
XX 09-AUG-2001.  
PD  
XX 30-JAN-2001; 2001WO-US000654.  
PF  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-48898/53.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
PT  
XX Claim 27; SEQ ID NO 37262; 658pp; English.  
PS  
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 34 AA;  
SQ  
Query Match 1.8%; Score 6; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 287 TQPIVE 292  
Db 1 TQPIVE 6  
RESULT 198  
ABG46059  
ID ABG46059 standard; peptide; 34 AA.  
XX  
AC ABG46059;  
XX  
XX 19-AUG-2002 (first entry)  
DT  
XX Human peptide encoded by genome-derived single exon probe SEQ ID 35724.  
DE  
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
KW

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
XX WO200186003-A2.  
FN  
XX 15-NOV-2001.  
PD  
XX 30-JAN-2001; 2001WO-US000665.  
PF  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2002-114183/15.  
DR  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
PT  
XX Claim 27; SEQ ID NO 35724; 634pp; English.  
PS  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC; the novel set of probes which hybridise at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a peptide/protein encoded by a single exon probe of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

```

XX SQ Sequence 34 AA;
Query Match 1.8%; Score 6; DB 5; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TQPIVE 292
Db 1 TQPIVE 6

RESULT 199
AAR63517
ID AAR63517 standard; peptide; 38 AA.
XX AC AAR63517;
XX DT 25-MAR-2003 (revised)
XX AC 22-JUN-1995 (first entry)
XX DE Bovine neutrophil beta-defensin peptide BNBED-8.
XX KW Cattle; bovine; neutrophil; beta-defensin; antibiotic; antimicrobial;
XX KW pharmaceutical.
XX OS Bos taurus.
XX FN W09421672-A1.
XX PD 29-SEP-1994.
XX PF 15-MAR-1994; 94WO-US003178.
XX PR 19-MAR-1993; 93US-00033873.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Selsted ME, Cullor JS;
XX DR WPI; 1994-316934/39.
XX PT Purified beta-defensin peptide(s) - useful as antimicrobial agents in
XX PT human and veterinary medicine.
XX PS Disclosure; Fig 5; 79pp; English.
XX CC The peptide is useful in a pharmaceutical composition as an antimicrobial
XX CC compound for use in human and veterinary medicine and as an agent in
XX CC agricultural, food science and industrial applications. (Updated on 25-
XX CC MAR-2003 to correct PN field.)
XX SQ Sequence 38 AA;
Query Match 1.8%; Score 6; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RRQIGT 111
Db 23 RRQIGT 28

RESULT 200
ADE03026
ID ADE03026 standard; peptide; 39 AA.
XX AC ADE03026;
XX DT 29-JAN-2004 (first entry)
XX DE Hybrid polypeptide pharmacokinetic enhancer peptide, SEQ ID No 1533.
XX XX

XX SQ Sequence 39 AA;
Query Match 1.8%; Score 6; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQE 72
Db 15 AQLAQE 20

RESULT 201
ADB67852
ID ADB67852 standard; protein; 39 AA.
XX AC ADB67852;
XX DT 04-DEC-2003 (first entry)
XX DE Human lung specific protein sequence SEQ ID NO:135.
XX XX

XX hybrid; enhancer; anti-fusogenic; antiviral; virucide; antidiabetic;
XX pharmacokinetic; fusogenic; insulin; diabetes.
XX Unidentified.
XX Key Location/Qualifiers
XX Modified-site 1
XX Modified-site 39 /note= "Residue is modified by acetyl group"
XX Modified-site 39 /note= "C-terminal amide"
XX US6348568-B1.
XX 19-FEB-2002.
XX 20-MAY-1999; 99US-00315304.
XX 20-MAY-1998; 98US-00082279.
XX (TRIM-) TRIMERIS INC.
XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX WPI; 2002-424396/45.
XX New hybrid polypeptide for modulating fusogenic events for e.g. antiviral
XX activity, has enhancer peptide sequence derived from retroviral envelope
XX protein sequences linked to core polypeptide e.g. therapeutic protein.
XX Disclosure; SEQ ID NO 1533; 70pp; English.
XX The invention relates to a novel hybrid polypeptide comprising an
XX enhancer peptide sequence linked to a core polypeptide. The enhancer
XX peptide sequence comprises WQSEQKI or WASMEWF. The invention also
XX includes novel peptides that exhibit anti-fusogenic activity, antiviral
XX activity and/or ability to modulate intracellular processes. The novel
XX hybrid polypeptide has virucide and antidiabetic activity. The enhancer
XX peptide sequence enhances pharmacokinetic properties of any core
XX polypeptide, for example, a polypeptide useful for the treatment or
XX prevention of a disease, or an imaging agent useful for imaging
XX structures in vivo. The core polypeptides and hybrid polypeptides are
XX useful for modulating fusogenic events and exhibit antifusogenic or
XX antiviral activity. The novel hybrid polypeptide is useful for decreasing
XX viral infection and modulating intracellular processes involving coiled-
XX coil peptide interactions. The novel hybrid polypeptide comprises insulin
XX or its fragment, so the core polypeptide is useful for ameliorating the
XX symptoms of forms of diabetes. The novel hybrid polypeptide is also
XX useful as a part of prognosis for preventing disorders including fusion
XX events and viral infection that involves cell-cell and/or virus-cell
XX fusion, and for diagnosis and in vivo imaging methods. This sequence
XX represents an enhancer peptide of the invention.
XX SQ Sequence 39 AA;

```



KW human; lung specific polypeptide; LSP; lung specific nucleic acid; LSNA;  
 KW lung; cytostatic; vaccine; gene therapy; immune response; lung cancer;  
 XX metastasis.

OS Homo sapiens.

PN WO2003020899-A2.

XX 13-MAR-2003.

XX 29-AUG-2002; 2002WO-US027771.

XX 31-AUG-2001; 2001US-0316260P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Liu C, Chen S;

XX WPI; 2003-300880/29.

XX N-PSDB; ADB67946.

XX Novel lung specific polypeptides and nucleic acids, useful for

PT identifying, diagnosing, monitoring, staging, imaging and treating lung

PT cancer and non-cancerous disease states in lung.

XX Example 1; Page 243; 263pp; English.

CC The present invention describes human lung specific polypeptides (LSP,  
 CC (I)) and lung specific nucleic acids (LSNA, (II)). Also described: (1)  
 CC determining (MI) the presence of an LSNA in a sample; (2) a vector (III)  
 CC comprising (II); (3) a host cell comprising (III); (4) a polypeptide (IV)  
 CC encoded by (II); (5) production of (IV); (6) an antibody (V) or its  
 CC fragment that specifically binds to an LSP; (7) determining (M2) the  
 CC presence of an LSP in a sample; (8) a kit for detecting a risk of cancer  
 CC or presence of cancer in a patient, comprising a unit for determining the  
 CC presence of (I) or (II) in a sample of the patient; and (9) a vaccine  
 CC comprising (I) or (II). LSP and LSNA sequences have cytostatic  
 CC activities, and can be used in vaccines, gene therapy, and as inducers of  
 CC an immune response. (I) and (II) can be used for diagnosing or monitoring  
 CC the presence and metastases of lung cancer in a patient, by determining  
 CC the amount of (I) or (II) in a sample of a patient, and comparing the  
 CC amount of the determined nucleic acid molecule or the polypeptide in the  
 CC sample of the patient to the amount of the lung specific marker in a  
 CC normal control, where the difference in the amount of the nucleic acid  
 CC molecule or the polypeptide in the sample compared to the amount of the  
 CC nucleic acid molecule or the polypeptide in the normal control is  
 CC associated with the presence of lung cancer. (I) and (V) are useful for  
 CC treating a patient with lung cancer, by administering (I) or (V), where  
 CC the administration of (I) or (V) induces an immune response against the  
 CC lung cancer cell expressing the nucleic acid molecule or a polypeptide.  
 CC (I), (II) and (V) are also useful for identifying, diagnosing,  
 CC monitoring, staging, imaging and treating lung cancer and non-cancerous  
 CC disease states in lung, for identifying lung tissue, and monitoring,  
 CC identifying and/or designing agonists and antagonists of (I). (II) can be  
 CC used in gene therapy, for producing transgenic animals and cells, and for  
 CC producing engineered lung tissue for treatment and research. (II) is also  
 CC useful for detecting or amplifying nucleic acid molecules that have  
 CC similar or identical nucleic acid sequences compared to (II). (I) is  
 CC useful for producing engineered lung tissue. (V) is useful for  
 CC identifying (II). (II) is also useful for driving in vivo expression of  
 CC (I). The present sequence represents a human LSP from the present  
 CC invention.

XX Sequence 39 AA;

Query Match 1.8%; Score 6; DB 7; Length 39;

Best Local Similarity 100.0%; Pred. No. 6.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 STFDIA 180

DB 19 STFDIA 24

RESULT 202

AAR63516

ID AAR63516 standard; peptide; 40 AA.

XX AC AAR63516;

XX DT 25-MAR-2003 (revised)

XX DT 22-JUN-1995 (first entry)

XX DE Bovine neutrophil beta-defensin peptide BNBD-7.

XX KW Cattle; bovine; neutrophil; beta-defensin; antibiotic; antimicrobial;

XX KW pharmaceutical.

XX OS Bos taurus.

XX PN WO9421672-A1.

XX PD 29-SEP-1994.

XX PF 15-MAR-1994; 94WO-US003178.

XX PR 19-MAR-1993; 93US-00033873.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Selsted ME, Cullor JS;

XX DR WPI; 1994-316934/39.

XX PT Purified beta-defensin peptide(s) - useful as antimicrobial agents in

XX human and veterinary medicine.

XX PS Disclosure; Fig 5; 79pp; English.

XX CC The peptide is useful in a pharmaceutical composition as an antimicrobial

XX compound for use in human and veterinary medicine and as an agent in

XX agricultural, food science and industrial applications. (Updated on 25-

XX MAR-2003 to correct PN field.)

XX SQ Sequence 40 AA;

Query Match 1.8%; Score 6; DB 2; Length 40;

Best Local Similarity 100.0%; Pred. No. 6.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111

DB 25 RQIGT 30

RESULT 203

AAR63518

ID AAR63518 standard; peptide; 40 AA.

XX AC AAR63518;

XX DT 25-MAR-2003 (revised)

XX DT 22-JUN-1995 (first entry)

XX DE Bovine neutrophil beta-defensin peptide BNBD-9.

XX KW Cattle; bovine; neutrophil; beta-defensin; antibiotic; antimicrobial;

XX KW pharmaceutical.

XX OS Bos taurus.

XX PN WO9421672-A1.

XX PD 29-SEP-1994.

XX PF 15-MAR-1994; 94WO-US003178.

XX 19-MAR-1993; 93US-00033873.  
 XX (REGC ) UNIV CALIFORNIA.  
 XX Selsted ME, Cullor JS;  
 XX WPI; 1994-316934/39.  
 XX Purified beta-defensin peptide(s) - useful as antimicrobial agents in  
 PT human and veterinary medicine.  
 XX Disclosure; Fig 5; 79pp; English.  
 XX The peptide is useful in a pharmaceutical composition as an antimicrobial  
 CC compound for use in human and veterinary medicine and as an agent in  
 CC agricultural, food science and industrial applications. (Updated on 25-  
 CC MAR-2003 to correct PN field.)  
 XX SQ Sequence 40 AA;  
 Query Match 1.8%; Score 6; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 106 RQIGT 111  
 Db 25 RQIGT 30  
 RESULT 204  
 AAU91027  
 ID AAU91027 standard; peptide; 40 AA.  
 AC AAU91027;  
 XX 05-JUN-2002 (first entry)  
 XX Transplant media associated defensin peptide #28.  
 XX Transplant; antimicrobial peptide; pore forming agent;  
 KW cell surface receptor binding compound; kidney transplant; cardioplegia;  
 KW organ transplant; transplant rejection; defensin.  
 XX Bos taurus.  
 XX WO200209738-A1.  
 XX 07-FEB-2002.  
 XX 27-JUL-2001; 2001WO-US023785.  
 XX 28-JUL-2000; 2000US-0221632P.  
 XX 17-NOV-2000; 2000US-0249602P.  
 XX 15-MAY-2001; 2001US-0290932P.  
 XX (MURP/) MURPHY C J.  
 XX Murphy CJ, Reid TW, Meanulty JF;  
 XX WPI; 2002-268995/31.  
 XX Media comprising antimicrobial polypeptides or pore forming agents and/or  
 PT cell surface receptor binding compounds useful for the storage and  
 PT preservation of organs prior to transplant.  
 XX Claim 8; Page 29; 78pp; English.  
 XX The invention describes new transplant compositions comprising  
 CC antimicrobial polypeptides or pore forming agents and/or cell surface  
 CC receptor binding compounds. The media is capable of extending the  
 CC preservation period past 72 hours and can provide organs with increased  
 CC functionality upon transplant. animals receiving kidneys stored in the  
 CC media of the present invention for either three or four days had serum  
 CC creatinine levels of less than half of those observed in control animals  
 CC receiving kidneys stored in UW solution (defined in the specification)  
 CC alone. Lower serum creatinine levels are indicative of healthier kidneys  
 CC and a more preferable prognosis for the transplant patient. The media of  
 CC the invention are useful for decreasing the incidence and/or severity of

CC media of the present invention for either three or four days had serum  
 CC creatinine levels of less than half of those observed in control animals  
 CC receiving kidneys stored in UW solution (defined in the specification)  
 CC alone. Lower serum creatinine levels are indicative of healthier kidneys  
 CC and a more preferable prognosis for the transplant patient. The media of  
 CC the invention are useful for decreasing the incidence and/or severity of  
 CC delayed graft function in patients receiving transplanted kidneys stored  
 CC and/or treated in the media. The media may also be used in procedures  
 CC such as cardioplegia. It is contemplated that transplant of healthier  
 CC organs leads to a decrease in chronic rejection. This sequence represents  
 CC an antimicrobial defensin peptide studied in the development of the  
 XX transplant media  
 XX SQ Sequence 40 AA;  
 Query Match 1.8%; Score 6; DB 5; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 106 RQIGT 111  
 Db 25 RQIGT 30  
 RESULT 205  
 AAU91026  
 ID AAU91026 standard; peptide; 40 AA.  
 AC AAU91026;  
 XX 05-JUN-2002 (first entry)  
 XX Transplant media associated defensin peptide #27.  
 XX Transplant; antimicrobial peptide; pore forming agent;  
 KW cell surface receptor binding compound; kidney transplant; cardioplegia;  
 KW organ transplant; transplant rejection; defensin.  
 XX Bos taurus.  
 XX WO200209738-A1.  
 XX 07-FEB-2002.  
 XX 27-JUL-2001; 2001WO-US023785.  
 XX 28-JUL-2000; 2000US-0221632P.  
 XX 17-NOV-2000; 2000US-0249602P.  
 XX 15-MAY-2001; 2001US-0290932P.  
 XX (MURP/) MURPHY C J.  
 XX Murphy CJ, Reid TW, Meanulty JF;  
 XX WPI; 2002-268995/31.  
 XX Media comprising antimicrobial polypeptides or pore forming agents and/or  
 PT cell surface receptor binding compounds useful for the storage and  
 PT preservation of organs prior to transplant.  
 XX Claim 8; Page 29; 78pp; English.  
 XX The invention describes new transplant compositions comprising  
 CC antimicrobial polypeptides or pore forming agents and/or cell surface  
 CC receptor binding compounds. The media is capable of extending the  
 CC preservation period past 72 hours and can provide organs with increased  
 CC functionality upon transplant. animals receiving kidneys stored in the  
 CC media of the present invention for either three or four days had serum  
 CC creatinine levels of less than half of those observed in control animals  
 CC receiving kidneys stored in UW solution (defined in the specification)  
 CC alone. Lower serum creatinine levels are indicative of healthier kidneys  
 CC and a more preferable prognosis for the transplant patient. The media of  
 CC the invention are useful for decreasing the incidence and/or severity of

CC delayed graft function in patients receiving transplanted kidneys stored  
 CC and/or treated in the media. The media may also be used in procedures  
 CC such as cardioplegia. It is contemplated that transplant of healthier  
 CC organs leads to a decrease in chronic rejection. This sequence represents  
 CC an antimicrobial defensin peptide studied in the development of the  
 CC transplant media  
 XX  
 XX Sequence 40 AA;  
 SQ

Query Match 1.8%; Score 6; DB 5; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RRRQIGT 111  
 DB 25 RRRQIGT 30  
 |||||

RESULT 206  
 ADA89676  
 ID ADA89676 standard; protein; 41 AA.  
 XX  
 AC ADA89676;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Staphylococcus aureus antigenic protein #215.  
 XX  
 KW antigenic protein; vaccine; Staphylococcus aureus; pathogenic organism;  
 KW antibacterial; neuroprotective; immunosuppressive; antiinflammatory;  
 KW antiulcer; immunostimulant; ophthalmological; pathogenic microbe;  
 KW bacteraemia; septic shock; organ infection; skin infection;  
 KW bacterial basal colonisation; bacterial eye infection; septicemia;  
 KW tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;  
 KW sepsis; meningitis; pneumonia; stomach ulcer; gonorrhoea; toxic shock;  
 KW necrotising fasciitis; impetigo; histoplasmosis; Lyme disease;  
 KW Gastro-enteritis; dysentery; shigellosis; skin disorder.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 XX WO2003011899-A2.  
 PN  
 XX 13-FEB-2003.  
 PD  
 XX 02-AUG-2002; 2002WO-GB003606.  
 PF  
 XX 02-AUG-2001; 2001GB-00018825.  
 PR  
 XX 03-JAN-2002; 2002GB-00000349.  
 XX  
 (UYSH-) UNIV SHEFFIELD.  
 PA (BIOS-) BIOSYNEXUS INC.  
 PA  
 XX Foster S, Mond J, Clarke S, McDowell P, Brummel K;  
 PI  
 XX WPI; 2003-256434/25.  
 DR  
 XX New antigenic polypeptides from Staphylococcus aureus or S. epidermidis,  
 PT useful as a vaccine for immunizing humans against e.g. bacteremia, septic  
 PT shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or  
 PT impetigo.  
 XX  
 Claim 4; Page 163; 189pp; English.  
 PS  
 XX The present invention describes an antigenic protein or its part, which  
 CC is for use as a vaccine. The antigenic protein is encoded by an isolated  
 CC DNA molecule of Staphylococcus aureus or S. epidermidis partial gene  
 CC sequences (designated dnasa and dna SE, respectively; and which encodes a  
 CC protein expressed by a pathogenic organism. Also described: (1) a vaccine  
 CC composition comprising at least one antigenic protein; (2) a method of  
 CC immunising an animal against a disease or condition caused by a  
 CC pathogenic microbe by administering the antigenic protein or the vaccine;  
 CC (3) an antibody or its binding part obtainable by the method above; (4)  
 CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a

CC hybridoma cell line produced by the method of (4); and (6) identifying  
 CC opsonic antigens expressed by a pathogenic microbe. The antigenic  
 CC proteins have antibacterial, neuroprotective, immunosuppressive,  
 CC antiinflammatory, antiulcer, immunostimulant and ophthalmological  
 CC activities, and can be used in vaccines. The antigenic proteins or  
 CC vaccines can be used for immunising an animal (specifically a human)  
 CC against a disease or condition caused by a pathogenic microbe, e.g.  
 CC bacteraemia, septic shock, organ infection, skin infection, bacterial  
 CC basal colonisation, bacterial eye infections, septicemia, tuberculosis,  
 CC bacteria-associated food poisoning, blood infections, peritonitis,  
 CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,  
 CC strep throat, streptococcal-associated toxic shock, necrotising  
 CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,  
 CC dysentery, shigellosis, S. aureus-associated septicemia, food-poisoning,  
 CC skin disorders, S. epidermidis-associated septicemia, peritonitis or  
 CC endocarditis. The present sequence represents a S. aureus antigenic  
 CC protein sequence from the present invention.  
 XX  
 SQ Sequence 41 AA;

Query Match 1.8%; Score 6; DB 6; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 LVADFL 202  
 DB 36 LVADFL 41  
 |||||

RESULT 207  
 ABB38318  
 ID ABB38318 standard; peptide; 42 AA.  
 XX  
 AC ABB38318;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 DE Peptide #5824 encoded by human foetal liver single exon probe.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US000669.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 WPI; 2001-483447/52.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human fetal liver.  
 XX  
 PS Claim 27; SEQ ID NO 30953; 639pp + Sequence Listing, English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe



XX Peptide #557 encoded by probe for measuring cervical gene expression.  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX Homo sapiens.  
 OS WO200157278-A2.  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000670.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488901/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 PT  
 XX Claim 27; SEQ ID NO 18949; 487pp; English.  
 PS The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see AAI10069-AAL28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 43 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0;  
 QY 77 GLLVTL 82  
 Db 35 GLLVTL 40  
 RESULT 211  
 ABB33068  
 ID ABB33068 standard; peptide; 43 AA.  
 XX AC ABB33068;  
 XX DT 06-FEB-2002 (first entry)  
 XX Peptide #574 encoded by human foetal liver single exon probe.  
 DE Human; foetal liver; gene expression; single exon nucleic acid probe.  
 KW Homo sapiens.  
 OS WO200157277-A2.  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000663.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-483447/52.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human fetal liver.  
 PT  
 XX Claim 27; SEQ ID NO 25703; 639pp + Sequence Listing; English.  
 PS The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 43 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0;  
 QY 77 GLLVTL 82  
 Db 35 GLLVTL 40  
 RESULT 212  
 AAM26530  
 ID AAM26530 standard; protein; 43 AA.  
 XX AC AAM26530;  
 XX DT 18-OCT-2001 (first entry)  
 XX Peptide #567 encoded by probe for measuring placental gene expression.  
 DE Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX Homo sapiens.  
 OS WO200157272-A2.  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000663.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI



PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226682P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 06-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 12-SEP-2000; 2000US-0232081P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 25-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 26-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 13-OCT-2000; 2000US-0239337P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246509P.

PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 05-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 05-JAN-2001; 2001US-0254097P.  
 XX  
 (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

WFI; 2001-465570/50.

N-PSDB; AAL00940.

Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.

Claim 11; SEQ ID NO 3628; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a protein of the invention

Sequence 43 AA;

Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 FLSFQ 309

Db 27 FLSFQ 32

RESULT 215

ABBI8537

ID ABBI8537 standard; protein; 43 AA.

XX ABBI8537;

AC ABBI8537;

DT 24-JAN-2002 (first entry)

XX Protein #536 encoded by probe for measuring heart cell gene expression.

DE



XX Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.

XX Homo sapiens.

PN WO200157274-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488999/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts.

XX Claim 15; SEQ ID NO 20307; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

XX measuring human gene expression in a sample derived from human heart (see

XX AB21535-ABA41305). The present sequence is a protein encoded by one such

XX probe. The probes may be used for predicting, measuring and displaying

XX gene expression in samples derived from the human heart via microarrays.

XX By measuring gene expression, the probes are useful for predicting,

XX diagnosing, grading, staging, monitoring and prognosing diseases of the

XX human heart and vascular system e.g. cardiovascular disease,

XX hypertension, cardiac arrhythmias and congenital heart disease. Note: The

XX sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 43 AA;

XX Query Match 1.8%; Score 6; DB 4; Length 43;

XX Best Local Similarity 100.0%; Pred. No. 7.1e+02; Indels 0; Gaps 0;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 77 GLLVTL 82

XX Db 35 GLLVTL 40

XX RESULT 216

XX AAM66253

XX ID AAM66253 standard; protein; 43 AA.

XX AC AAM66253;

XX XX 06-NOV-2001 (first entry)

XX XX Human bone marrow expressed probe encoded protein SEQ ID NO: 26559.

XX XX Human; bone marrow expressed exon; gene expression analysis; probe;

XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX XX WO200157276-A2.

XX PN Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX (MOLE-) MOLECULAR DYNAMICS INC.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488999/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

XX Example 4; SEQ ID NO 26559; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention

XX Sequence 43 AA;

XX Query Match 1.8%; Score 6; DB 4; Length 43;

XX Best Local Similarity 100.0%; Pred. No. 7.1e+02; Indels 0; Gaps 0;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 77 GLLVTL 82

XX Db 35 GLLVTL 40

XX RESULT 217

XX AAM53865

XX ID AAM53865 standard; protein; 43 AA.

XX AC AAM53865;

XX XX 05-NOV-2001 (first entry)

XX XX Human brain expressed single exon probe encoded protein SEQ ID NO: 25970.

XX XX Human; brain expressed exon; gene expression analysis; probe; microarray;

XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

XX OS Homo sapiens.

XX XX WO200157275-A2.

XX PN Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX (MOLE-) MOLECULAR DYNAMICS INC.

XX XX 30-JAN-2001; 2001WO-US000667.

XX XX 04-FEB-2000; 2000US-0180312P.

XX XX 26-MAY-2000; 2000US-0207456P.

XX XX 30-JUN-2000; 2000US-00608408.

XX XX 03-AUG-2000; 2000US-00632366.

XX XX 21-SEP-2000; 2000US-0234687P.

XX XX 27-SEP-2000; 2000US-0236359P.

XX XX 04-OCT-2000; 2000GB-00024263.

XX XX (MOLE-) MOLECULAR DYNAMICS INC.

XX XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX (MOLE-) MOLECULAR DYNAMICS INC.

XX WPI; 2001-483446/52.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.  
 XX Example 4; SEQ ID NO 25970; 650pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention  
 XX  
 XX Sequence 43 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 77 GLLVTL 82  
 Db 35 GLLVTL 40  
 |||||  
 RESULT 218  
 ABG47918  
 ID ABG47918 standard; peptide; 43 AA.  
 AC ABG47918;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human liver peptide, SEQ ID NO 26566.  
 XX  
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157273-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488898/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 XX  
 XX Claim 27; SEQ ID NO 26566; 658pp; English.  
 XX  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression

CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
 CC liver single exon encoded peptides of the invention. Note: The sequence  
 CC information for this patent does not appear in the printed specification  
 CC but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 43 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 77 GLLVTL 82  
 Db 35 GLLVTL 40  
 |||||  
 RESULT 219  
 ABB95680  
 ID ABB95680 standard; protein; 43 AA.  
 XX  
 AC ABB95680;  
 XX  
 DT 21-JUN-2002 (first entry)  
 XX  
 DE Human testicular antigen SEQ ID NO: 1064.  
 XX  
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
 KW reproductive system disorder; urinary system disorder; gene therapy;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disease; infection; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155317-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001329.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228944P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232388P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 23-SEP-2000; 2000US-0234997P.  
PR 23-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 28-SEP-2000; 2000US-0236327P.  
PR 28-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 13-OCT-2000; 2000US-0239379P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
PT for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Claim 11; SEQ ID NO 1064; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a protein of the  
CC invention  
XX  
XX Sequence 43 AA;  
SQ

Query Match 1.8%; Score 6; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 304 FLSFQ 309  
DB 27 FLSFQ 32

RESULT 220  
AA001861  
ID AA001861 standard; protein; 43 AA.  
XX  
XX AA001861;  
XX  
XX DT 09-OCT-2001 (first entry)  
XX  
XX Peptide #543 encoded by probe for measuring human breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX Homo sapiens.  
 OS  
 XX WO200157270-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX  
 XX 29-JAN-2001; 2001WO-US000661.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR  
 XX 26-MAY-2000; 2000US-0207456P.  
 PR  
 XX 30-JUN-2000; 2000US-00608408.  
 PR  
 XX 03-AUG-2000; 2000US-00632366.  
 PR  
 XX 21-SEP-2000; 2000US-0234687P.  
 PR  
 XX 27-SEP-2000; 2000US-0236359P.  
 PR  
 XX 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-476286/51.  
 DR  
 XX Novel single exon nucleic acid probe used to measuring gene expression in  
 PT a human breast.  
 PT  
 XX Claim 27; SEQ ID NO 10601; 322pp; English.  
 PS  
 XX The present invention relates to novel single exon nucleic acid probes  
 CC (see AAL00010-AAL10067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in  
 CC a human breast sample, where the probe hybridises at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 43 AA;  
 SQ  
 Query Match 1.8%; Score 6; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 77 GLLVTL 82  
 DB 35 GLLVTL 40  
 |||||  
 RESULT 221  
 ABG35900  
 ID ABG35900 standard; peptide; 43 AA.  
 XX  
 AC ABG35900;  
 XX  
 XX 20-AUG-2002 (first entry)  
 DT  
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 25565.  
 DE  
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX Homo sapiens.  
 OS  
 XX WO200186003-A2.  
 PN  
 XX 15-NOV-2001.  
 PD  
 XX  
 XX 30-JAN-2001; 2001WO-US000665.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR  
 XX 26-MAY-2000; 2000US-0207456P.  
 PR  
 XX 30-JUN-2000; 2000US-00608408.  
 PR  
 XX 03-AUG-2000; 2000US-00632366.  
 PR  
 XX 21-SEP-2000; 2000US-0234687P.  
 PR  
 XX 27-SEP-2000; 2000US-0236359P.  
 PR  
 XX 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2002-114183/15.  
 DR  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 PT  
 XX Claim 27; SEQ ID NO 25565; 634pp; English.  
 PS  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a nucleic  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a peptide/protein encoded by a single exon probe of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 43 AA;  
 Query Match 1.8%; Score 6; DB 5; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 77 GLLVTL 82
Db 35 GLLVTL 40
RESULT 222
AAG18732
ID AAG18732 standard; protein; 44 AA.
AC AAG18732;
XX
DT 17-OCT-2000 (first entry)
DE Zea mays protein fragment SEQ ID NO: 20258.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
FD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
```

PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 26-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154773P.  
 PR 22-SEP-1999; 99US-0155133P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155653P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match  
 Best Local Similarity 100.0%; DB 3; Length 44;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 YSSGLL 79  
 Db 5 YSSGLL 10

RESULT 223  
 AAUI4265  
 ID AAUI4265 standard; protein; 44 AA.  
 AC AAUI4265;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human novel protein #136.  
 XX  
 KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antithrombotic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US002623.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-451939/48.  
 DR N-PSDB; AAS22570.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 XX nervous system disorders, and for regenerating bone and cartilage.  
 PS Example 4; Page 600; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human proteins or  
 CC their active domains. The polypeptides, polynucleotides and antibodies  
 CC raised against the polypeptides are used in a method of treatment of a  
 CC mammal and prevention of disorders caused by the aberrant protein  
 CC expression or activity. The polypeptides can be used as molecular weight  
 CC markers, food supplements, and in antibody production. The polypeptides  
 CC are used to identify compounds which bind to the polypeptides.  
 CC Polynucleotides of the invention are used as probes and primers, for  
 CC sequencing, for chromosome or gene mapping, in the production of  
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 CC therapy. Polypeptides of the invention can be used to target drugs to a  
 CC tumour, in assays to determine biological activity, to raise  
 CC antibodies/elicite an immune response, to determine quantitative protein  
 CC levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory  
 CC diseases, nervous system disorders, and infection. The present sequence  
 CC represents a protein of the invention  
 XX  
 SQ Sequence 44 AA;

Query Match  
 Best Local Similarity 100.0%; DB 4; Length 44;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 KIILFS 162

Db 36 KILFS 41  
|||||

## RESULT 224

AAO04708  
ID AAO04708 standard; protein; 45 AA.

XX AAO04708;

AC AAO04708;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 18600.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

PN 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;  
PI WPI; 2001-514838/56.

XX N-PSDB; AA184639.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
and treating e.g. leukemia, inflammation and immune disorders.

Claim 20; SEQ ID NO 18600; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and  
the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
cytokine, cell proliferation or cell differentiation or which may induce  
production of other cytokines in other cell populations. The  
polynucleotides and polypeptides are useful in gene therapy, vaccines or  
e.g. stem cell growth factor activity, haematopoiesis regulating  
activity, tissue growth factor activity, immunomodulatory activity and  
activin/inhibin activity and may be useful in the diagnosis and/or  
treatment of cancer, leukaemia, nervous system disorders, arthritis and  
inflammation. Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic format  
directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 45 AA;

## Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233

Db 19 SLKLLG 24

## RESULT 225

ABG09887

ID ABG09887 standard; protein; 45 AA.

XX ABG09887;

AC ABG09887;

DT 20-MAR-2003 (revised)

14-MAY-1999 (first entry)

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #9878.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

XX N-PSDB; AA574074.

New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.

Claim 20; SEQ ID NO 40246; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II)  
sequences. (I) is useful as hybridisation probes, polymerase chain  
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
and in recombinant production of (II). The polynucleotides are also used  
in diagnostics as expressed sequence tags for identifying expressed  
genes. (I) is useful in gene therapy techniques to restore normal  
activity of (II) or to treat disease states involving (II). (II) is  
useful for generating antibodies against it, detecting or quantitating a  
polypeptide in tissue, as molecular weight markers and as a food  
supplement. (II) and its binding partners are useful in medical imaging  
of sites expressing (II). (I) and (II) are useful for treating disorders  
involving aberrant protein expression or biological activities in  
polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
amino acid sequences of the invention. Note: The sequence data for this  
patent did not appear in the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 45 AA;

## Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LFSKSH 11

Db 23 LFSKSH 28

## RESULT 226

AAW92823

ID AAW92823 standard; protein; 46 AA.

XX AAW92823;

AC AAW92823;

DT 20-MAR-2003 (revised)

14-MAY-1999 (first entry)



```

XX DE Human transaldolase TAL-H protein fragment #6.
XX KW Transaldolase; TAL-H; autoantigen; human; antibody; neuroprotector;
XX KW neurodegenerative autoimmune disease; multiple sclerosis; detection;
XX KW proliferation inhibitor; cytotoxicity inhibitor; binding inhibitor.
XX OS Homo sapiens.
XX XX US5879909-A.
XX PD 09-MAR-1999.
XX PF 09-APR-1998; 98US-00057762.
XX PR 19-OCT-1994; 94US-00326119.
XX PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX PI Perl A;
XX DR WPI; 1999-203948/17.
XX PT Isolated human transaldolase gene - useful for raising antibodies for
XX PT detecting neurodegenerative autoimmune diseases, especially multiple
XX PT sclerosis.
XX PS Claim 5; Col 59-60; 55pp; English.
XX CC This sequence represents a human transaldolase (TAL-H) protein fragment.
XX CC This protein fragment can be used in methods to raise antibodies for
XX CC detecting human transaldolase-mediated neurodegenerative autoimmune
XX CC diseases, especially multiple sclerosis. The protein has neuroprotective
XX CC properties and is a proliferation, cytotoxicity and binding inhibitor.
XX CC (Updated on 20-MAR-2003 to correct PF field.)
XX SQ Sequence 46 AA;
XX Query Match 1.8%; Score 6; DB 2; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 7.6e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 230 KLLGEL 235
XX DB 15 KLLGEL 20
XX RESULT 227
XX ID AAY68326 standard; peptide; 46 AA.
XX AC AAY68326;
XX XX 18-APR-2000 (first entry)
XX DT Human transaldolase protein (TAL-H) peptide SEQ ID NO:21.
XX DE Human; transaldolase; autoantigen; multiple sclerosis; diagnosis;
XX KW neurodegenerative disease; neuroprotective; TAL-H.
XX KW Homo sapiens.
XX OS US6018021-A.
XX PN 25-JAN-2000.
XX PD 19-OCT-1994; 94US-00326119.
XX PF 19-OCT-1994; 94US-00326119.
XX PR (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX PA Perl A;
XX PI

XX DR WPI; 2000-136702/12.
XX PT Transaldolase proteins and peptides, useful for diagnosing multiple
XX PT sclerosis.
XX PS Claim 4; Col 11-12; 55pp; English.
XX CC The present invention describes a human transaldolase protein molecule
XX CC (TAL-H). Peptides (I) derived from TAL-H which comprise at least 1 r cell
XX CC epitope or at least 1 B cell/antibody epitope are also described. TAL-H
XX CC or (I) induce a transaldolase-specific immune response. TAL-H, (I) and
XX CC antibodies to TAL-H or (I) are useful for diagnosing and treating
XX CC multiple sclerosis and other human immune-related neurodegenerative
XX CC diseases. The present sequence represents a specifically example of a TAL
XX CC -H peptide, from the present invention
XX SQ Sequence 46 AA;
XX Query Match 1.8%; Score 6; DB 3; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 7.6e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 230 KLLGEL 235
XX DB 15 KLLGEL 20
XX RESULT 228
XX ID AAM92813
XX AC AAM92813 standard; protein; 47 AA.
XX AC AAM92813;
XX DT 06-NOV-2001 (first entry)
XX DE Human digestive system antigen SEQ ID NO: 2162.
XX KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX KW digestive system disorder; Meckel's diverticulum.
XX OS Homo sapiens.
XX XX WO200155314-A2.
XX PN 02-AUG-2001.
XX PD 17-JAN-2001; 2001WO-US001324.
XX PF 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.

```

PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225457P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228242P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 08-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 21-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 05-JAN-2001; 2001US-0254097P.  
 XX 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-502630/55.  
 XX N-PSDB; AAK88586.

PT Polynucleotides encoding digestive system antigens, useful for  
 PT diagnosing, treating, preventing and/or prognosing disorders of the  
 PT digestive system, particularly cancer and cancer metastases.

XX Claim 11; SEQ ID NO 2162; 986pp; English.

XX The present invention provides the protein and coding sequences of a  
 CC number of human digestive system antigens. These can be used in the  
 CC diagnosis, treatment and prevention of digestive system disorders,  
 CC including cancer, Meckel's diverticulum, bacterial or parasitic  
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
 CC ulcerative colitis. The present sequence is a digestive system antigen of  
 XX the invention

SQ Sequence 47 AA;

Query Match

Best Local Similarity 100.0%; Score 6; DB 4; Length 47;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107

Db 20 NNILRR 25

## RESULT 229

AAO05232  
ID AAO05232 standard; protein; 47 AA.  
AC AAO05232;

XX  
DT 06-NOV-2001 (first entry)

XX  
DE Human polypeptide SEQ ID NO 19124.

XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX  
OS Homo sapiens.

XX  
FN WO200164835-A2.

XX  
PD 07-SEP-2001.

XX  
PF 26-FEB-2001; 2001WO-US004927.

XX  
PR 28-FEB-2000; 2000US-00515126.

XX  
PR 18-MAY-2000; 2000US-00577409.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PI Tang YT, Liu C, Drmanac RT;

XX  
DR WPI; 2001-514838/56.

XX  
DR N-PSDB; AAI85163.

XX  
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
XX and treating e.g. leukemia, inflammation and immune disorders.

XX  
PS Claim 20; SEQ ID NO 19124; 1399pp + Sequence Listing; English.

XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 47 AA;

Query Match 1.8%; Score 6; DB 4; Length 47;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLG 233

Db 3 SLKLG 8

## RESULT 230

AAU22668  
ID AAU22668 standard; protein; 47 AA.

XX  
AC AAU22668;

XX  
DT 18-DEC-2001 (first entry)

XX  
DE Novel human colon associated polypeptide #201.

KW Human; colon cancer; congenital abnormality; infection; colitis;  
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;  
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;  
KW sigmoid disease; antibacterial; antiviral; antiinflammatory; cytostatic.  
OS Homo sapiens.  
XX  
FN WO200155302-A2.  
XX  
PD 02-AUG-2001.

XX  
PF 17-JAN-2001; 2001WO-US001240.

XX  
PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 23-AUG-2000; 2000US-0227182P.

PR 30-AUG-2000; 2000US-0227009P.

PR 01-SEP-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 05-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 08-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 14-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465567/50.  
DR N-PSDB; AAS39548.  
XX  
PT Isolated polypeptide for treating, preventing and/or prognosing  
disorders related to the colon including colon cancers and also for  
testing and detection e.g. diagnosis.

XX  
PS Claim 11; SEQ ID NO 445; 562pp; English.  
XX  
CC The present invention relates to the isolation of novel human colon  
associated polypeptides, and the cDNA (AAS39348-AAS39581) and genomic  
sequences encoding for them. The sequences of the invention are useful in  
the diagnosis, treatment, prevention and/or prognosis of disorders of the  
colon including colon cancer, congenital abnormalities (e.g. atresia and  
stenosis), bacterial and viral infections, inflammatory bowel disease  
(IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal  
inflammatory disorders, colitis, colonic inflammation, diarrhoea and  
dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal  
obstruction and sigmoid diseases. The polynucleotide sequences of the  
invention can also be used in gene therapy. AAU2468-AAU2701 represent  
the novel human colon associated polypeptides of the invention. Note: The  
sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO

XX SQ Sequence 47 AA;

Query Match 1.8%; Score 6; DB 4; Length 47;  
Best Local Similarity 100.0%; Pred.No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107  
Db 20 NNILRR 25

RESULT 231  
ADB32508  
ID ADB32508 standard; protein; 47 AA.  
XX  
AC ADB32508;  
XX  
DT 04-DEC-2003 (first entry)  
DE Human novel colon related polypeptide SEQ ID NO 445.  
XX  
KW gene therapy; cancer; liver disorder; hepatitis; neural disorder;  
Alzheimer's disease; human; colon.  
OS Homo sapiens.  
XX  
PN US2003050231-Al.  
PD 13-MAR-2003.  
XX  
PF 17-JAN-2001; 2001US-00764872.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0225113P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0244647P.  
 PR 08-NOV-2000; 2000US-0244647P.  
 PR 08-NOV-2000; 2000US-0244647P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 (ROSE/) ROSEN C A.  
 (RUBE/) RUBEN S M.  
 (BARA/) BARASH S C.  
 Rosen CA, Ruben SM, Barash SC;  
 WPI; 2003-625420/59.  
 N-PSDB; ADB32274.  
 New nucleic acid molecule, useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g. cancer, liver disorders or neural disorders.  
 Claim 12; SEQ ID NO 445; 216pp; English.  
 The invention relates to an isolated nucleic acid molecule encoding a polypeptide. The nucleic acid is useful for preparing a medicament for

CC preventing, treating or ameliorating a medical condition e.g. cancer,  
 CC liver disorders such as hepatitis or neural disorders such as Alzheimer's  
 CC disease. The present sequence represents the amino acid sequence of a  
 CC human novel colon related polypeptide. Note: The sequence data for this  
 CC patent did not form part of the printed specification but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20030050231.

XX Sequence 47 AA;

Query Match 1.8%; Score 6; DB 7; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107

Db 20 NNILRR 25

RESULT 232

ID ADE78262 standard; protein; 48 AA.

XX ADE78262;

XX 29-JAN-2004 (first entry)

XX Endometrial specific protein identified as DEX0379\_20\_aa\_1 (SeqID 169).  
 XX neoplastic; endometrial cell; endometrial specific nucleic acid; ESNA;  
 XX cancer; metastasis; gene therapy; cytostatic; human.

XX Homo sapiens.

XX WO2003055982-A2.

XX 10-JUL-2003.

XX 20-DEC-2002; 2002WO-US041175.

XX 21-DEC-2001; 2001US-0343134P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Liu C;

XX WPI; 2003-569441/53.

XX N-PSDB; ADE78367.

XX New nucleic acid, useful for preparing a composition for diagnosing or  
 XX treating endometrial cancer.

XX Claim 1; SEQ ID NO 169; 327pp; English.

XX This invention relates to novel nucleic acid molecules, and encoded  
 XX proteins thereof, which are present in normal and neoplastic endometrial  
 XX cells. Specifically, it refers to the use of these endometrial specific  
 XX nucleic acids (ESNAs), as well as suitable antibodies, agonists and  
 XX antagonists that are useful for the identification, diagnosis and  
 XX monitoring of endometrial cancer. The present invention describes a  
 XX method for monitoring the presence of an endometrial specific protein in  
 XX a sample (potentially representing metastases) that comprises contacting  
 XX the sample with a reagent and detecting the intensity of this  
 XX interaction. Accordingly, via gene therapy, these ESNAs are useful for  
 XX preparing a cytostatic composition for diagnosing or treating endometrial  
 XX cancer. This polypeptide sequence is a human endometrial specific  
 XX polypeptide of the invention.

XX Sequence 48 AA;

Query Match 1.8%; Score 6; DB 7; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLV 80  
 Db 35 SSGLLV 40

RESULT 233

AAB34514  
 ID AAB34514 standard; protein; 49 AA.

XX AAB34514;

XX 02-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 132.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
 XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 XX neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO200056767-A1.

XX 28-SEP-2000.

XX 16-MAR-2000; 2000WO-US006828.

XX 19-MAR-1999; 99US-0125358P.

XX 08-DEC-1999; 99US-0169616P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-602216/57.

XX Nucleic acid molecules encoding human secreted proteins, used in  
 XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 XX Parkinson's diseases and cancers.

XX Disclosure; Page 368; 384pp; English.

XX The invention relates to the isolation of genes AAC59624-CS9669 encoding  
 XX the human secreted proteins AAB34439-B34484. This sequence represents a  
 XX peptide fragment homologous to the protein encoded by the gene isolated  
 XX in the present invention. The sequence is used as a query sequence for  
 XX doing BLASTX searches to determine homologous sequence to the protein.  
 XX The genes and proteins are useful for preventing, ameliorating or  
 XX treating medical conditions, e.g. by protein or gene therapy. The genes  
 XX are isolated from a range of human tissues disclosed in the  
 XX specification. The nucleic acids, proteins, antibodies and (ant)agonists  
 XX are useful in the diagnosis, treatment and prevention of: (a) cancer,  
 XX e.g. breast and ovarian cancer, and other cancers of the adrenal gland,  
 XX bone, bone marrow, breast, gastrointestinal tract, liver, lung, or  
 XX urogenital; (b) immune disorders e.g. Addison's disease, allergies,  
 XX autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
 XX Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 XX colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)  
 XX wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 XX and parasitic infections

XX Sequence 49 AA;

Query Match 1.8%; Score 6; DB 3; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233

|||||

Db 18 SLKLLG 23

RESULT 234  
AAG75342  
ID AAG75342 standard; protein; 50 AA.

XX  
AC AAG75342;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE Human colon cancer antigen protein SEQ ID NO:6106.  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN WO200122920-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-US026524.  
XX  
PR 29-SEP-1999; 99US-0157137P.  
PR 03-NOV-1999; 99US-0163280P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI; 2001-235357/24.  
DR N-PSDB; AAH34747.  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers.  
XX  
PS Claim 11; Page 7564; 9803pp; English.

XX  
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patients own production of P. Additionally, N may be used  
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
XX to 1052, 7921 and 7922  
XX  
SQ Sequence 50 AA;

Query Match 1.8%; Score 6; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEAVAQ 68  
Db 15 TEAVAQ 20

RESULT 235  
AAU45565  
ID AAU45565 standard; protein; 50 AA.  
XX

AC AAU45565;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #6461.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.  
DR N-PSDB; AAS59526.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
PS Example 1; SEQ ID NO 6760; 1069pp; English.

XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 50 AA;

Query Match 1.8%; Score 6; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LLQSEN 221  
Db 21 LLQSEN 26

RESULT 236  
AAU42299  
ID AAU42299 standard; protein; 50 AA.  
XX



AC AAU42299;  
XX 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #3195.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2003-616774/71.  
DR N-PSDB; AAS59516.  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
XX Example 1; SEQ ID NO 3494; 1069pp; English.  
PS  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 50 AA;  
Query Match 1.8%; Score 6; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 LLVTLI 83  
Db 2 LLVTLI 7  
RESULT 237  
ABM42084  
ID ABM42084 standard; protein; 50 AA.  
XX

AC ABM42084;  
XX 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #6760.  
XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032727.  
XX  
PR 15-OCT-2001; 2001US-00978825.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallieue-Douglas J;  
XX  
XX WPI; 2003-381789/36.  
DR N-PSDB; ACF64455.  
XX  
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
XX Example 1; SEQ ID NO 6760; 1481pp; English.  
PS  
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 50 AA;

Query Match 1.8%; Score 6; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 216 LLQSEN 221  
Db 21 LLQSEN 26

RESULT 238  
ID ABM38818 standard; protein; 50 AA.  
XX  
AC ABM38818;  
XX  
DT 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #3494.  
XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032727.  
XX  
PR 15-OCT-2001; 2001US-00978825.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Valliave-Douglass J;  
XX  
DR WPI: 2003-381789/36.  
DR N-PSDB; ACF64445.  
XX  
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
PS Example 1; SEQ ID NO 3494; 1481pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 50 AA;  
Query Match 1.8%; Score 6; DB 6; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 78 LLVTLI 83

RESULT 239  
ID AAO04380 standard; protein; 51 AA.  
XX  
AC AAO04380;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 18272.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US004927.  
XX  
PR 28-FEB-2000; 2000US-00515126.  
PR 18-MAY-2000; 2000US-00577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
DR N-PSDB; AAI84311.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
PS Claim 20; SEQ ID NO 18272; 1399pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 51 AA;  
Query Match 1.8%; Score 6; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 228 SLKLLG 233  
DB 39 SLKLLG 44

RESULT 240  
ID AAB34389 standard; protein; 52 AA.  
XX  
AC AAB34389;  
XX

DT 26-JAN-2001 (first entry)  
DE Human secreted protein sequence encoded by gene 23 SEQ ID NO:150.  
XX  
KW Human; secreted protein; diagnosis; neuroprotective; cytoskeletal;  
KW cardioactive; immunomodulatory; muscular active general; vulnery;  
KW gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW and antibacterial; gene therapy; detection; cancer; chromosome marker;  
KW chromosome identification; neural disorder; immune disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; wound healing; infectious disease; preservative;  
KW food additive.  
XX  
OS Homo sapiens.  
XX  
XX WO200056883-A1.  
XX  
PD 28-SEP-2000.  
XX  
XX 16-MAR-2000; 2000WO-US006822.  
XX  
PF 23-MAR-1999; 99US-0126054P.  
XX  
PR 10-DEC-1999; 99US-0169916P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-587666/55.  
XX  
XX Human secreted proteins and gene sequences encoding them, useful for  
XX detecting, preventing, and treating disorders such as cancer,  
XX neurological disorders and immune system disorders.  
XX  
PS Disclosure; Page 406; 429pp; English.  
XX  
XX The polynucleotide sequences given in AAC59566 to AAC59614 encode the  
XX human secreted proteins given in AAB34299 to AAB34347. AAB34348 to  
XX AAB34347 represent human secreted polypeptide sequences and proteins  
XX homologous to them, which are given in the exemplification of the present  
XX invention. Human secreted proteins have activities based on the tissues  
XX and cells the genes are expressed in. Example of activities include:  
XX neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular  
XX active general; vulnery; gastrointestinal; nephrotropic; antiinfective;  
XX gynaecological; and antibacterial. The polynucleotides can be used for  
XX the detection of various disorders such as cancer, chromosome  
XX identification, as chromosome markers, and for numerous other diagnostic  
XX or research purposes. The secreted proteins can be used to treat  
XX disorders such as neural, immune, muscular, reproductive,  
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
XX disorders, wound healing, and infectious diseases. The proteins can also  
XX be used as a food additive or preservative to increase or decrease  
XX storage capabilities. AAC59567 to AAC59565 and AAB34298 represent  
XX sequences used in the exemplification of the present invention  
XX  
SQ Sequence 52 AA;  
Query Match 1.8%; Score 6; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 228 SLKLLG 233  
Db 22 SLKLLG 27  
|||||  
RESULT 241  
AAU53123  
ID AAU53123 standard; protein; 53 AA.  
XX  
XX AAU53123;  
AC  
XX

DT 27-FEB-2002 (first entry)  
DE Propionibacterium acnes immunogenic protein #14019.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
XX WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
XX 20-APR-2001; 2001WO-US012865.  
XX  
XX 21-APR-2000; 2000US-0199047P.  
XX  
PR 02-JUN-2000; 2000US-020841P.  
XX  
PR 07-JUL-2000; 2000US-021674P.  
XX  
XX (CORI-), CORIXA CORP.  
XX  
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
XX WPI; 2001-616774/71.  
XX  
DR N-PSDB; AAS59558.  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
XX vaccinating against and diagnosing infections, especially useful for  
XX treating acne vulgaris.  
XX  
PS Example 1; SEQ ID NO 14318; 1069pp; English.  
XX  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
XX polypeptides. The proteins and their associated DNA sequences are used in  
XX the treatment, prevention and diagnosis of medical conditions caused by  
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
XX P. acnes is also involved in infections of bone, joints and the central  
XX nervous system, however it is particularly involved in the inflammatory  
XX lesions associated with acne vulgaris. A method for detecting the  
XX presence or absence of P. acnes in a patient comprises contacting a  
XX sample with a binding agent that binds to the proteins of the invention  
XX and determining the amount of bound protein in the sample. The  
XX polypeptides may be used as antigens in the production of antibodies  
XX specific for P. acnes proteins. These antibodies can be used to  
XX downregulate expression and activity of P. acnes polypeptides and  
XX therefore treat P. acnes infections. The antibodies may also be used as  
XX diagnostic agents for determining P. acnes presence, for example, by  
XX enzyme linked immunosorbent assay (ELISA). Note: the sequence data for  
XX this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 53 AA;  
Query Match 1.8%; Score 6; DB 4; Length 53;  
Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 136 PQIALR 141  
Db 42 PQIALR 47  
|||||  
RESULT 242  
ABM49642  
ID ABM49642 standard; protein; 53 AA.  
XX  
XX ABM49642;  
AC  
XX

20-OCT-2003 (first entry)  
 Propionibacterium acnes predicted ORF-encoded polypeptide #14318.  
 Acne vulgaris; antisborrheic; dermatological; antibacterial;  
 immunostimulant; immune response; vaccine.  
 Propionibacterium acnes.  
 WO2003033515-A1.  
 24-APR-2003.  
 11-OCT-2002; 2002WO-US032727.  
 15-OCT-2001; 2001US-00978825.  
 (CORI-) CORIXA CORP.  
 Mittham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
 Barth B, Valliere-Douglass J;  
 WPI; 2003-381789/36.  
 N-PSDB; ACF64487.  
 New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 or for stimulating an immune response specific for a P. acnes protein.  
 Example 1; SEQ ID NO 14318; 1481pp; English.  
 The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 encoding a Propionibacterium acnes protein. The invention also relates to  
 polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 immunogenic fragments of P. acnes polypeptides. The invention  
 additionally encompasses expression vectors and host cells comprising a  
 polynucleotide of the invention; antibodies against polypeptides of the  
 invention; fusion proteins comprising a polypeptide of the invention; a  
 method for stimulating an immune response specific for a P. acnes  
 polypeptide and an isolated T cell population comprising T cells prepared  
 via this method; a vaccine composition (comprising P. acnes polypeptides,  
 polynucleotides, antibodies, fusion proteins, T cell populations, or  
 antigen-presenting cells that express the polypeptide); a method and kit  
 for detecting or determining the presence or absence of P. acnes in a  
 patient; and a method for inhibiting the presence or absence of P. acnes in a  
 patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 proteins, T cell populations or antigen-presenting cells that express the  
 polypeptides are useful for diagnosing, preventing or treating acne  
 vulgaris, or for stimulating an immune response specific for a P. acnes  
 protein. The polynucleotides can also be used as probes or primers for  
 nucleic acid hybridisation. The vaccine composition is useful for the  
 stimulation of an immune response against P. acnes, or for treating acne,  
 and the kit is useful for performing a diagnostic assay. The present  
 sequence represents a polypeptide predicted to be encoded by an ORF (open  
 reading frame) contained within the P. acnes polynucleotide of the  
 invention. Note: The sequence data for this patent did not form part of  
 the printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Query Match 1.8%; Score 6; DB 6; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 136 PQIALR 141  
 DB 42 PQIALR 47  
 RESULT 243  
 AA012979

AA012979 standard; protein; 54 AA.  
 AA012979;  
 06-NOV-2001 (first entry)  
 Human polypeptide SEQ ID NO 26871.  
 Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 tissue growth factor; immunomodulatory; cancer; leukaemia;  
 nervous system disorders; arthritis; inflammation.  
 Homo sapiens.  
 WO200164835-A2.  
 07-SEP-2001.  
 26-FEB-2001; 2001WO-US004927.  
 28-FEB-2000; 2000US-00515126.  
 18-MAY-2000; 2000US-00577409.  
 (HYSE-) HYSEQ INC.  
 Tang YT, Liu C, Drmanac RT;  
 WPI; 2001-514838/56.  
 N-PSDB; AA192910.  
 Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 and treating e.g. leukemia, inflammation and immune disorders.  
 Claim 20; SEQ ID NO 26871; 1399pp + Sequence Listing; English.  
 The invention relates to human polynucleotides (AA179941-AA193841) and  
 the encoded proteins (AA00010-AA03910) that exhibit activity relating to  
 cytokine, cell proliferation or cell differentiation or which may induce  
 production of other cytokines in other cell populations. The  
 polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 peptide therapy. The polypeptides have various cytokine-like activities,  
 e.g. stem cell growth factor activity, haematopoiesis regulating  
 activity, tissue growth factor activity, immunomodulatory activity and  
 activin/inhibin activity and may be useful in the diagnosis and/or  
 treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 inflammation. Note: The sequence data for this patent did not form part  
 of the printed specification, but was obtained in electronic format  
 directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Query Match 1.8%; Score 6; DB 4; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 228 SLKILG 233  
 DB 36 SLKILG 41  
 RESULT 244  
 AA047338  
 ID AAU47338 standard; protein; 54 AA.  
 AC AAU47338;  
 XX  
 DT 27-FEB-2002 (first entry)  
 DE Propionibacterium acnes immunogenic protein #8234.  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
XX 01-NOV-2001.  
XX  
XX 20-APR-2001; 2001WO-US012865.  
XX  
XX 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
XX WPI; 2001-616774/71.  
DR N-PSDB; AAS59538.  
XX  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
XX Example 1; SEQ ID NO 8533; 1069pp; English.  
XX  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 54 AA;  
  
Query Match 1.8%; Score 6; DB 4; Length 54;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 146 LRECIR 151  
Db 12 LRECIR 17  
|||||  
RESULT 245  
ABP07511  
ID ABP07511 standard; protein; 54 AA.  
XX  
AC ABP07511;  
XX  
XX 24-JUN-2002 (first entry)  
XX  
XX Human ORFX protein sequence SEQ ID NO:15004.  
DE  
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.  
XX  
OS Homo sapiens.  
XX  
PN WO200192523-A2.  
XX  
XX 06-DEC-2001.  
XX  
XX 29-MAY-2001; 2001WO-US010836.  
XX  
XX 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Leach MD;  
PI  
XX WPI; 2002-106308/14.  
DR N-PSDB; ABN23263.  
XX  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX  
XX Disclosure; SEQ ID NO 15004; 1037pp; English.  
XX  
XX The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 54 AA;  
  
Query Match 1.8%; Score 6; DB 5; Length 54;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 155 LAKIIL 160  
Db 47 LAKIIL 52  
|||||  
RESULT 246  
ABM43857  
ID ABM43857 standard; protein; 54 AA.  
XX  
XX ABM43857;  
AC  
XX

DT 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #8533.  
XX  
KW Acne vulgaris; antisborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
XX 11-OCT-2002; 2002WO-US032727.  
PF  
XX 15-OCT-2001; 2001US-00978825.  
PR  
XX (CORI-) CORIXA CORP.  
PA  
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallie-Douglass J;  
XX  
DR WPI; 2003-381789/36.  
DR N-PSDB; ACF64467.  
XX  
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
XX Example 1; SEQ ID NO 8533; 1481pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the presence or treating the  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 54 AA;  
  
Query Match 1.8%; Score 6; DB 6; Length 54;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 146 LRECIR 151  
Db 12 LRECIR 17  
  
RESULT 247  
AAG11542

ID AAG11542 standard; protein; 55 AA.  
XX  
AC AAG11542;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10297.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 07-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0134941P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.

```
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145219P.
PR 23-JUL-1999; 99US-0145219P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145911P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 1.8%; Score 6; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 9e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 LSSEFK 310
Db 19 LSSEFK 24

RESULT 248
ABP01202
ID ABP01202 standard; protein; 55 AA.
XX
AC ABP01202;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:2386.
```



XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.  
XX Homo sapiens.  
XX WO200192523-A2.  
XX 06-DEC-2001.  
XX 29-MAY-2001; 2001WO-US010836.  
XX 30-MAY-2000; 2000US-0206132P.  
XX 29-AUG-2000; 2000US-0228716P.  
XX (CURA-) CURAGEN CORP.  
XX Shinkets RA, Leach MD;  
XX WPI; 2002-106308/14.  
XX N-PSDB; ABN16954.  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX Disclosure; SEQ ID NO 2386; 1037pp; English.  
XX The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 55 AA;  
Query Match 1.8%; Score 6; DB 5; Length 55;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 128 MLLKGY 133  
DB 1 MLLKGY 6  
RESULT 249  
ABP25762  
ID ABP25762 standard; protein; 55 AA.

XX ABP25762;  
XX 02-JUL-2002 (first entry)  
XX Streptococcus polypeptide SEQ ID NO 700.  
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antinflammatory; infection; vaccine; meningitis; gene therapy.  
XX Streptococcus agalactiae.  
XX WO200234771-A2.  
XX 02-MAY-2002.  
XX 29-OCT-2001; 2001WO-GB004789.  
XX 27-OCT-2000; 2000GB-00026333.  
XX 24-NOV-2000; 2000GB-00028727.  
XX 07-MAR-2001; 2001GB-00005640.  
XX (CHIR-) CHIRON SPA.  
XX (GENO-) INST GENOMIC RES.  
XX Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;  
XX Tettelin H;  
XX WPI; 2002-352536/38.  
XX N-PSDB; ABN66393.  
XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.  
XX Claim 1; Page 3222; 4525pp; English.  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins  
XX Sequence 55 AA;  
Query Match 1.8%; Score 6; DB 5; Length 55;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 93 GKQDVT 98  
DB 24 GKQDVT 29  
RESULT 250  
AAM94423  
ID AAM94423 standard; protein; 56 AA.  
XX AAM94423;  
XX

DT 21-NOV-2001 (first entry)  
XX Human reproductive system related antigen SEQ ID NO: 3081.  
DE Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy.  
KW Homo sapiens.  
XX  
OS  
XX  
XX WO20015320-A2.  
PN  
XX  
XX 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001339.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 27-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232089P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465570/50.  
 DR N-PSDB; AAL00393.  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
 PT used in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Claim 11; SEQ ID NO 3081; 1297pp + Sequence Listing; English.  
 XX  
 CC The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a protein of the invention  
 XX  
 SQ Sequence 56 AA;  
 Query Match 1.8%; Score 6; DB 4; Length 56;  
 Best Local Similarity 100.0%; Pred.No. 9.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 VKILKD 23  
 Db |||||  
 6 VKILKD 11

Search completed: April 12, 2004, 10:35:23  
 Job time : 75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:36:54 ; Search time 46 Seconds  
(without alignments)  
1926.306 Million cell updates/sec

Title: US-10-025-730-1

Perfect score: 337

Sequence: 1 MKOMPLFSKHNPAEIVKI.....FADEKNYLIKQIRDLKKTAP 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1073127 seqs, 262937947 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	337	100.0	337	14	US-10-025-730-1
2	23	6.8	341	14	US-10-025-730-3
3	23	6.8	496	14	US-10-239-079-5
4	23	6.8	552	14	US-10-239-079-6
5	15	4.5	339	14	US-10-025-730-4
6	14	4.2	377	14	US-10-025-730-5
7	10	3.0	342	12	US-10-425-114-48872
8	9	2.7	339	12	US-10-424-599-224185
9	8	2.4	184	12	US-10-424-599-149595
10	8	2.4	307	12	US-10-369-493-8871
11	8	2.4	412	12	US-10-425-114-54669
12	8	2.4	446	12	US-10-425-114-521177
13	8	2.4	492	12	US-10-425-114-55961
14	8	2.4	549	15	US-10-369-493-13726
15	8	2.4	923	9	US-09-862-027-23
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 5, Appli
					Sequence 48872, A
					Sequence 224185,
					Sequence 149595,
					Sequence 8871, Ap
					Sequence 54669, A
					Sequence 52177, A
					Sequence 55961, A
					Sequence 13726, A
					Sequence 23, Appl

16 7 2.1 51 14 US-10-238-075-1453 Sequence 1453, Ap

17 7 2.1 63 12 US-10-424-599-230655 Sequence 230655, A

18 7 2.1 83 12 US-10-425-114-44450 Sequence 44450, A

19 7 2.1 87 12 US-10-424-599-165543 Sequence 165543, A

20 7 2.1 95 12 US-10-424-599-230733 Sequence 230733, A

21 7 2.1 99 9 US-09-731-872-430 Sequence 430, App

22 7 2.1 99 10 US-09-876-997-430 Sequence 430, App

23 7 2.1 102 12 US-10-424-599-194139 Sequence 194139, A

24 7 2.1 113 14 US-10-106-698-5213 Sequence 5213, Ap

25 7 2.1 133 14 US-10-156-761-8084 Sequence 8084, Ap

26 7 2.1 141 12 US-10-282-122A-59346 Sequence 59346, A

27 7 2.1 147 12 US-10-424-599-222167 Sequence 222167, A

28 7 2.1 151 12 US-10-425-114-56304 Sequence 56304, A

29 7 2.1 157 12 US-10-424-599-234035 Sequence 234035, A

30 7 2.1 161 12 US-10-425-114-70029 Sequence 70029, A

31 7 2.1 165 12 US-10-424-599-256726 Sequence 256726, A

32 7 2.1 168 12 US-10-424-599-173625 Sequence 173625, A

33 7 2.1 168 12 US-10-424-599-221069 Sequence 221069, A

34 7 2.1 168 12 US-10-425-114-38325 Sequence 38325, A

35 7 2.1 173 12 US-10-424-599-158029 Sequence 158029, A

36 7 2.1 176 9 US-09-729-674-280 Sequence 280, App

37 7 2.1 197 15 US-10-108-260A-4437 Sequence 4437, Ap

38 7 2.1 199 15 US-10-108-260A-4032 Sequence 4032, Ap

39 7 2.1 204 12 US-10-425-114-45029 Sequence 45029, A

40 7 2.1 215 10 US-09-791-279-185 Sequence 185, App

41 7 2.1 217 15 US-10-094-749-2522 Sequence 2522, Ap

42 7 2.1 227 14 US-10-156-761-8653 Sequence 8653, Ap

43 7 2.1 243 12 US-10-424-599-203478 Sequence 203478, A

44 7 2.1 245 12 US-10-424-599-179206 Sequence 179206, A

45 7 2.1 245 14 US-10-005-549-17 Sequence 17, Appl

46 7 2.1 251 12 US-10-424-599-151467 Sequence 151467, A

47 7 2.1 275 14 US-10-080-170-328 Sequence 328, App

48 7 2.1 282 12 US-10-282-122A-52103 Sequence 52103, A

49 7 2.1 337 12 US-10-424-599-209662 Sequence 209662, A

50 7 2.1 338 12 US-10-435-114-72787 Sequence 72787, A

51 7 2.1 341 12 US-10-424-599-234484 Sequence 234484, A

52 7 2.1 344 12 US-10-282-122A-51574 Sequence 51574, A

53 7 2.1 351 12 US-10-425-114-46025 Sequence 46025, A

54 7 2.1 356 12 US-10-424-599-262896 Sequence 262896, A

55 7 2.1 360 12 US-10-425-114-41971 Sequence 41971, A

56 7 2.1 360 12 US-10-425-114-47953 Sequence 47953, A

57 7 2.1 373 15 US-10-369-493-42 Sequence 42, Appl

58 7 2.1 374 14 US-10-106-698-4458 Sequence 4458, Ap

59 7 2.1 391 12 US-10-425-114-68619 Sequence 68619, A

60 7 2.1 399 12 US-10-425-114-37084 Sequence 37084, A

61 7 2.1 422 14 US-10-005-549-2 Sequence 2, Appli

62 7 2.1 422 15 US-10-307-928A-2 Sequence 59712, A

63 7 2.1 439 12 US-10-282-122A-59712 Sequence 59712, A

64 7 2.1 448 12 US-10-425-114-37583 Sequence 37583, A

65 7 2.1 452 14 US-10-238-075-859 Sequence 859, App

66 7 2.1 458 14 US-10-032-585-7310 Sequence 7310, Ap

67 7 2.1 488 12 US-10-424-599-221067 Sequence 221067, A

68 7 2.1 490 12 US-09-952-464A-10 Sequence 10, Appl

69 7 2.1 490 12 US-10-425-114-38253 Sequence 38253, A

70 7 2.1 504 10 US-09-844-653-173 Sequence 173, App

71 7 2.1 504 10 US-09-852-463A-8 Sequence 8, Appli

72 7 2.1 504 13 US-10-086-135-4 Sequence 4, Appli

73 7 2.1 504 12 US-10-244-633-32 Sequence 32, Appl

74 7 2.1 508 12 US-10-425-114-55451 Sequence 55451, A

75 7 2.1 514 9 US-09-800-729-124 Sequence 124, App

76 7 2.1 514 11 US-09-833-245-2221 Sequence 2221, Ap

77 7 2.1 524 12 US-10-424-599-220386 Sequence 220386, A

78 7 2.1 534 16 US-10-389-566-822 Sequence 822, App

79 7 2.1 534 16 US-10-389-566-1163 Sequence 1163, Ap

80 7 2.1 558 15 US-10-369-493-14077 Sequence 14077, A

81 7 2.1 558 12 US-10-425-114-68937 Sequence 68937, A

82 7 2.1 571 15 US-10-407-866-93 Sequence 93, Appl

83 7 2.1 575 16 US-10-389-566-1261 Sequence 1261, Ap

84 7 2.1 575 16 US-10-389-566-2425 Sequence 2425, Ap

85 7 2.1 576 16 US-10-389-566-1939 Sequence 1939, Ap

86 7 2.1 578 12 US-10-425-114-66875 Sequence 66875, A

87 7 2.1 631 15 US-10-369-493-12179 Sequence 12179, A

88 7 2.1 637 12 US-10-425-114-64320 Sequence 64320, A

89	7	2.1	637	12	US-10-425-114-72475	Sequence 72475, A	162	74	12	US-10-424-599-244014	Sequence 244014,
90	7	2.1	731	13	US-10-086-464-8	Sequence 8, Appli	163	74	12	US-10-425-114-37263	Sequence 37263, A
91	7	2.1	762	15	US-10-369-293-2071	Sequence 2071, Ap	164	74	12	US-10-425-114-56749	Sequence 56749, A
92	7	2.1	855	15	US-10-312-088-34	Sequence 34, Appl	165	76	11	US-09-864-408A-2574	Sequence 2574, Ap
93	7	2.1	873	14	US-10-176-306-80	Sequence 30, Appl	166	77	9	US-09-867-550-574	Sequence 574, Ap
94	7	2.1	888	15	US-10-312-088-33	Sequence 33, Appl	167	78	12	US-10-424-599-253313	Sequence 253313,
95	7	2.1	891	12	US-10-282-122A-69986	Sequence 69986, A	168	78	13	US-10-002-344A-200	Sequence 200, App
96	7	2.1	893	15	US-10-314-143-27	Sequence 27, Appl	169	80	12	US-10-424-599-162228	Sequence 162228,
97	7	2.1	1053	12	US-10-282-122A-69548	Sequence 69548, A	170	80	12	US-10-424-599-198296	Sequence 198296,
98	7	2.1	1170	15	US-10-369-493-3006	Sequence 3006, Ap	171	80	12	US-10-424-599-244639	Sequence 244639,
99	7	2.1	1745	9	US-09-800-729-89	Sequence 89, Appl	172	81	9	US-09-738-626-4641	Sequence 4641, Ap
100	7	2.1	1762	11	US-09-833-245-2222	Sequence 222, Ap	173	82	9	US-09-858-664A-32	Sequence 32, Appl
101	7	2.1	1765	9	US-09-764-853-554	Sequence 2, Appli	174	82	9	US-09-764-877-1643	Sequence 1643, Ap
102	7	2.1	1766	9	US-10-399-385-5	Sequence 54, App	175	82	9	US-10-697-263-33	Sequence 33, Appl
103	6	1.8	11	12	US-10-399-385-5	Sequence 5, Appli	176	82	15	US-10-242-515-1643	Sequence 1643, Ap
104	6	1.8	14	14	US-10-261-798-51	Sequence 51, Appl	177	83	10	US-09-764-891-3412	Sequence 3412, Ap
105	6	1.8	15	12	US-10-399-385-4	Sequence 4, Appli	178	83	12	US-10-424-599-231338	Sequence 231338,
106	6	1.8	18	9	US-09-864-761-46460	Sequence 312, App	179	84	9	US-09-864-761-37566	Sequence 37566, A
107	6	1.8	24	10	US-09-774-639-312	Sequence 239, App	180	84	12	US-10-424-599-245645	Sequence 245645,
108	6	1.8	24	10	US-09-969-730-239	Sequence 239, App	181	85	12	US-10-424-599-186473	Sequence 186473,
109	6	1.8	24	16	US-10-621-363-239	Sequence 239, App	182	85	12	US-10-424-599-211413	Sequence 211413,
110	6	1.8	34	9	US-09-864-761-45850	Sequence 45850, A	183	86	9	US-09-925-301-1603	Sequence 1603, Ap
111	6	1.8	38	12	US-10-424-599-230826	Sequence 230826,	184	87	11	US-09-864-408A-8152	Sequence 8152, Ap
112	6	1.8	40	9	US-09-917-340-62	Sequence 62, Appl	185	87	12	US-10-424-599-153190	Sequence 153190,
113	6	1.8	40	9	US-09-917-340-63	Sequence 63, Appl	186	88	9	US-09-764-887-257	Sequence 257, App
114	6	1.8	42	9	US-09-864-761-46961	Sequence 46961, A	187	88	12	US-10-424-599-165968	Sequence 165968,
115	6	1.8	43	9	US-09-864-761-33835	Sequence 33835, A	188	88	12	US-10-424-599-251249	Sequence 251249,
116	6	1.8	43	10	US-09-764-891-3628	Sequence 3628, Ap	189	88	12	US-10-335-977-8935	Sequence 8935, Ap
117	6	1.8	44	15	US-10-291-265-372	Sequence 372, App	190	88	14	US-10-073-961-257	Sequence 257, App
118	6	1.8	45	12	US-10-424-599-157183	Sequence 157183,	191	89	12	US-10-424-599-258688	Sequence 258688,
119	6	1.8	46	12	US-10-424-599-159316	Sequence 159316,	192	89	12	US-10-335-977-8936	Sequence 8936, Ap
120	6	1.8	47	10	US-09-764-872-445	Sequence 445, App	193	90	12	US-09-882-227-114	Sequence 114, App
121	6	1.8	48	12	US-10-424-599-188181	Sequence 188181,	194	90	12	US-10-424-599-163560	Sequence 163560,
122	6	1.8	48	12	US-10-424-599-263667	Sequence 263667,	195	90	12	US-10-424-599-186702	Sequence 186702,
123	6	1.8	49	12	US-10-424-599-174437	Sequence 174437,	196	90	12	US-10-424-599-188262	Sequence 188262,
124	6	1.8	49	12	US-10-424-599-186215	Sequence 186215,	197	90	12	US-10-424-599-280571	Sequence 280571,
125	6	1.8	50	14	US-10-106-698-6116	Sequence 6116, Ap	198	91	12	US-10-424-599-206794	Sequence 206794,
126	6	1.8	51	12	US-10-424-599-223354	Sequence 223354,	199	92	11	US-09-864-408A-8758	Sequence 8758, Ap
127	6	1.8	52	12	US-10-424-599-249890	Sequence 249890,	200	92	12	US-10-424-599-157298	Sequence 157298,
128	6	1.8	52	14	US-10-156-761-15079	Sequence 15079, A	201	92	12	US-10-424-599-233814	Sequence 233814,
129	6	1.8	53	12	US-10-424-599-192699	Sequence 192699,	202	93	12	US-10-424-599-161289	Sequence 161289,
130	6	1.8	53	12	US-10-424-599-247013	Sequence 247013,	203	93	12	US-10-424-599-256649	Sequence 256649,
131	6	1.8	54	12	US-10-424-599-149472	Sequence 149472,	204	96	12	US-10-335-977-8937	Sequence 8937, Ap
132	6	1.8	55	12	US-10-424-599-182688	Sequence 182688,	205	97	12	US-10-424-599-234125	Sequence 234125,
133	6	1.8	56	9	US-09-908-711-83	Sequence 83, Appl	206	97	12	US-10-424-599-282678	Sequence 282678,
134	6	1.8	56	9	US-09-864-761-49022	Sequence 49022, A	207	97	12	US-10-424-599-284907	Sequence 284907,
135	6	1.8	56	10	US-09-764-891-3081	Sequence 3081, Ap	208	97	15	US-10-369-493-3935	Sequence 3935, Ap
136	6	1.8	56	12	US-10-424-599-267396	Sequence 267396,	209	99	12	US-10-424-599-147796	Sequence 147796,
137	6	1.8	57	12	US-10-424-599-178427	Sequence 178427,	210	99	12	US-10-424-599-155860	Sequence 155860,
138	6	1.8	57	14	US-10-029-386-27862	Sequence 27862, A	211	99	12	US-10-424-599-170968	Sequence 170968,
139	6	1.8	58	14	US-10-127-641-5	Sequence 5, Appli	212	99	12	US-10-424-599-194060	Sequence 194060,
140	6	1.8	60	12	US-10-424-599-195329	Sequence 195329,	213	100	12	US-10-424-599-164813	Sequence 164813,
141	6	1.8	62	12	US-10-424-599-259644	Sequence 259644,	214	101	12	US-10-424-599-146233	Sequence 146233,
142	6	1.8	62	12	US-10-424-599-191627	Sequence 191627,	215	101	15	US-10-094-749-1868	Sequence 1868, Ap
143	6	1.8	63	12	US-10-424-599-211257	Sequence 211257,	216	102	9	US-09-925-302-852	Sequence 852, App
144	6	1.8	63	12	US-10-424-599-209297	Sequence 209297,	217	102	9	US-09-864-761-3927	Sequence 3927, A
145	6	1.8	64	9	US-10-424-599-285560	Sequence 285560,	218	103	12	US-10-424-599-249236	Sequence 249236,
146	6	1.8	64	9	US-09-864-761-47397	Sequence 47397, A	219	104	9	US-09-867-550-76	Sequence 76, Appl
147	6	1.8	64	12	US-10-424-599-159823	Sequence 159823,	220	104	9	US-09-731-872-478	Sequence 478, App
148	6	1.8	64	12	US-10-335-977-5580	Sequence 5580, Ap	221	104	9	US-09-876-997-478	Sequence 478, App
149	6	1.8	65	12	US-10-424-599-215660	Sequence 215660,	222	105	9	US-09-771-161A-177	Sequence 177, App
150	6	1.8	66	12	US-10-424-599-165707	Sequence 165707,	223	105	12	US-10-424-599-216376	Sequence 216376,
151	6	1.8	67	12	US-10-424-599-191683	Sequence 191683,	224	106	14	US-10-263-828-83	Sequence 83, Appl
152	6	1.8	67	12	US-10-424-599-211264	Sequence 211264,	225	107	12	US-10-424-599-200671	Sequence 200671,
153	6	1.8	69	12	US-10-424-599-144751	Sequence 144751,	226	107	12	US-10-424-599-267548	Sequence 267548,
154	6	1.8	69	12	US-10-424-599-195946	Sequence 195946,	227	107	12	US-10-425-114-70895	Sequence 70895, A
155	6	1.8	69	12	US-10-424-599-202460	Sequence 202460,	228	108	12	US-10-424-599-178787	Sequence 178787,
156	6	1.8	69	12	US-10-424-599-262460	Sequence 262460,	229	108	12	US-10-424-599-191237	Sequence 191237,
157	6	1.8	71	12	US-10-221-278-242	Sequence 242, App	230	109	11	US-09-864-408A-6796	Sequence 6796, Ap
158	6	1.8	71	15	US-10-291-172-242	Sequence 242, App	231	109	12	US-10-282-122A-43428	Sequence 43428, A
159	6	1.8	72	12	US-10-282-122A-76423	Sequence 76423, A	232	109	12	US-10-282-122A-71923	Sequence 71923, A
160	6	1.8	73	11	US-09-864-408A-998	Sequence 998, App	233	109	12	US-10-424-599-258409	Sequence 258409,
161	6	1.8	73	12	US-10-424-599-203219	Sequence 203219,	234	110	12	US-10-425-114-46701	Sequence 46701, A

235 1.8 111 12 US-10-424-599-266911  
236 1.8 111 12 US-10-424-599-273101  
237 1.8 111 12 US-10-424-599-161030  
238 1.8 113 12 US-10-424-599-198662  
239 1.8 116 12 US-10-424-599-235299  
240 1.8 117 9 US-09-864-761-37175  
241 1.8 117 9 US-09-765-272-68  
242 1.8 117 9 US-09-939-980-347  
243 1.8 119 12 US-10-425-114-42612  
244 1.8 119 14 US-10-278-087A-50  
245 1.8 121 12 US-10-424-599-240972  
246 1.8 122 12 US-10-424-599-272118  
247 1.8 123 9 US-09-893-737-26  
248 1.8 125 12 US-10-424-599-175359  
249 1.8 125 12 US-10-424-599-186519  
250 1.8 125 12 US-10-424-599-243645  
251 1.8 126 9 US-09-864-761-48848  
252 1.8 126 15 US-10-294-006-38  
253 1.8 128 12 US-10-424-599-264673  
254 1.8 128 15 US-10-108-260A-3137  
255 1.8 128 15 US-10-108-260A-4117  
256 1.8 129 11 US-09-864-408A-2314  
257 1.8 133 12 US-10-282-122A-54727  
258 1.8 133 12 US-10-424-599-181434  
259 1.8 133 15 US-10-289-762-673  
260 1.8 134 15 US-10-080-334-30  
261 1.8 135 11 US-09-864-408A-6108  
262 1.8 135 12 US-10-424-599-284501  
263 1.8 136 10 US-09-809-391-554  
264 1.8 136 10 US-09-882-171-554  
265 1.8 136 12 US-10-425-114-66566  
266 1.8 136 15 US-10-080-334-26  
267 1.8 137 12 US-10-424-599-194879  
268 1.8 137 12 US-10-425-114-55006  
269 1.8 138 11 US-09-864-408A-1832  
270 1.8 138 12 US-10-424-599-273732  
271 1.8 139 12 US-10-424-599-178546  
272 1.8 140 9 US-09-764-869-987  
273 1.8 140 12 US-10-282-122A-71579  
274 1.8 140 12 US-10-424-599-199536  
275 1.8 140 12 US-10-424-599-223050  
276 1.8 140 14 US-10-091-504-987  
277 1.8 140 15 US-10-227-577-987  
278 1.8 142 12 US-10-282-122A-59233  
279 1.8 143 12 US-10-424-599-199815  
280 1.8 143 9 US-09-738-973-73  
281 1.8 144 9 US-09-854-133-73  
282 1.8 144 12 US-10-282-122A-78059  
283 1.8 144 12 US-10-424-599-155068  
284 1.8 144 12 US-10-424-599-180687  
285 1.8 144 14 US-10-144-649A-73  
286 1.8 145 12 US-10-282-122A-76902  
287 1.8 145 12 US-10-424-599-238754  
288 1.8 147 15 US-10-264-237-1997  
289 1.8 148 9 US-09-764-868-808  
290 1.8 148 11 US-09-764-875-1129  
291 1.8 148 14 US-10-170-102-8  
292 1.8 149 10 US-09-774-639-304  
293 1.8 149 10 US-09-969-730-231  
294 1.8 149 12 US-10-424-599-284490  
295 1.8 149 12 US-10-425-114-67031  
296 1.8 149 16 US-10-621-363-231  
297 1.8 152 12 US-10-282-132A-58999  
298 1.8 152 12 US-10-282-122A-78568  
299 1.8 152 12 US-10-335-977-8642  
300 1.8 152 14 US-10-291-851-89

ALIGNMENTS

; Sequence 1, Application US/10025730  
; Publication No. US20030045466A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/10/025,730  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: US/09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 3734805  
US-10-025-730-1

Query Match 100.0%; Score 337; DB 14; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKPLFKSHKNPAAIVKILKNDIAILEKQDKTKDSEVSKLQAMKEILCGTNEKE 60  
Db 1 MKKPLFKSHKNPAAIVKILKNDIAILEKQDKTKDSEVSKLQAMKEILCGTNEKE 60  
QY 61 PTEVAQAQLAQLYSSGLLVTLIAQLIDPEGKDVQIENNILRRQIGRSPVVEYS 120  
Db 61 PTEVAQAQLAQLYSSGLLVTLIAQLIDPEGKDVQIENNILRRQIGRSPVVEYS 120  
QY 121 AHPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSNQDFPFKYYVELSTFDIA 180  
Db 121 AHPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKILFNSNQDFPFKYYVELSTFDIA 180  
QY 181 SDAPATFKDILLTRHKVLVADFLQNYDIFEDYEKLLQSENVTYKQSLKGLGELILDRH 240  
Db 181 SDAPATFKDILLTRHKVLVADFLQNYDIFEDYEKLLQSENVTYKQSLKGLGELILDRH 240  
QY 241 NFAIMTKYISKPENLKMNLRLDKSPNIOPEAFHVEKVFVSPHKTQPIVEILLKNQPK 300  
Db 241 NFAIMTKYISKPENLKMNLRLDKSPNIOPEAFHVEKVFVSPHKTQPIVEILLKNQPK 300  
QY 301 LIEFLSSFKERTDDEQFADEKNYLIKQIRDLKKTAP 337  
Db 301 LIEFLSSFKERTDDEQFADEKNYLIKQIRDLKKTAP 337

RESULT 2  
US-10-025-730-3  
; Sequence 3, Application US/10025730  
; Publication No. US20030045466A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/10/025,730  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: US/09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Mus sp.

```
; FEATURE: -
; OTHER INFORMATION: g262934
US-10-025-730-3

Query Match          6.8%; Score 23; DB 14; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266
Db 245 IMTKYISKPENLKMNNLLRDKS 267

RESULT 3
US-10-239-079-5
; Sequence 5, Application US/10239079
; Publication No. US20030148446A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: -
; OTHER INFORMATION: Description of Artificial Sequence: Gal4-ANIC-BP-1
; OTHER INFORMATION: fusion protein
US-10-239-079-5

Query Match          6.8%; Score 23; DB 14; Length 496;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266
Db 400 IMTKYISKPENLKMNNLLRDKS 422

RESULT 4
US-10-239-079-6
; Sequence 6, Application US/10239079
; Publication No. US20030148446A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: -
; OTHER INFORMATION: Description of Artificial Sequence: LexA-ANIC-BP-1
; OTHER INFORMATION: fusion protein
US-10-239-079-6

Query Match          6.8%; Score 23; DB 14; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.8e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKMNNLLRDKS 266
Db 456 IMTKYISKPENLKMNNLLRDKS 478

RESULT 5
US-10-025-730-4
; Sequence 4, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: g1794137
US-10-025-730-4

Query Match          4.5%; Score 15; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPTEVEYI 119
Db 103 LRRQIGTRSPTEVEYI 117

RESULT 6
US-10-025-730-5
; Sequence 5, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-10-025-730-5

Query Match          4.2%; Score 14; DB 14; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPTEVEY 118
Db 119 LRRQIGTRSPTEVEY 132

RESULT 7
US-10-425-114-48872
; Sequence 48872, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
```



```
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48872
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3637-244-E1_FLI.pep
US-10-425-114-48872

Query Match          3.0%; Score 10; DB 12; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 AFHVKFVFA 282
DB 283 AFHVKFVFA 292
|||||

RESULT 8
US-10-424-599-224185
; Sequence 224185, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224185
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_44468C.1.pap
US-10-424-599-224185

Query Match          2.7%; Score 9; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 CGIMLRCCI 150
DB 146 CGIMLRCCI 154
|||||

RESULT 9
US-10-424-599-149595
; Sequence 149595, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
```

```
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149595
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106106C.1.pap
US-10-424-599-149595

Query Match          2.4%; Score 8; DB 12; Length 184;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 MLRECIRH 152
DB 140 MLRECIRH 147
|||||

RESULT 10
US-10-369-493-8871
; Sequence 8871, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8871
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Chloroflexus aurantiacus
US-10-369-493-8871

Query Match          2.4%; Score 8; DB 15; Length 307;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLIADL 86
DB 246 LVTLIADL 253
|||||

RESULT 11
US-10-425-114-54669
; Sequence 54669, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54669
; LENGTH: 412
; TYPE: PRT
```

; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3069-003-D7\_FLI.pep  
US-10-425-114-54669

Query Match 2.4%; Score 8; DB 12; Length 412;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 MLRECIRH 152  
DB 225 MLRECIRH 232

RESULT 12  
US-10-425-114-52177  
; Sequence 52177, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 52177  
; LENGTH: 446  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3912-007-H10\_FLI.pep  
US-10-425-114-52177

Query Match 2.4%; Score 8; DB 12; Length 446;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 MLRECIRH 152  
DB 259 MLRECIRH 266

RESULT 13  
US-10-425-114-55961  
; Sequence 55961, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 55961  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701125903\_FLI.pep  
US-10-425-114-55961

Query Match 2.4%; Score 8; DB 12; Length 492;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 DKASEEVS 43  
DB 202 DKASEEVS 209

RESULT 14  
US-10-369-493-13726  
; Sequence 13726, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Harry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 13726  
; LENGTH: 649  
; TYPE: PRT  
; ORGANISM: Pseudomonas fluorescens  
US-10-369-493-13726

Query Match 2.4%; Score 8; DB 15; Length 649;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 RQSLKLLG 233  
DB 381 RQSLKLLG 388

RESULT 15  
US-09-862-027-23  
; Sequence 23, Application US/09862027  
; Patent No. US20020142428A1  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin R.  
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof  
; FILE REFERENCE: 35800/234862  
; CURRENT APPLICATION NUMBER: US/09/862,027  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 09/345,473  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 923  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-862-027-23

Query Match 2.4%; Score 8; DB 9; Length 923;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLVTL 82  
DB 449 SSGLLVTL 456

RESULT 16  
US-10-238-075-1453

```
; Sequence 1453, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; FILE REFERENCE: E. coli, and biological uses of these polynucleotides and of their
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1453
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-1453

Query Match      2.1%  Score 7;  DB 14;  Length 51;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      292  EILLKQK 298
Db      13  EILLKQK 19

RESULT 17
US-10-424-599-230655
; Sequence 230655, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 230655
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_50302C.1.pep
US-10-424-599-230655

Query Match      2.1%  Score 7;  DB 12;  Length 63;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      272  EAFHVK 278
Db      4  EAFHVK 10

RESULT 18
US-10-425-114-44450
; Sequence 44450, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220733

Query Match      2.1%  Score 7;  DB 12;  Length 87;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      28  LEKQDKK 34
Db      24  LEKQDKK 30

RESULT 20
US-10-424-599-220733
; Sequence 220733, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220733

FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44450
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700959026_FLI.pep
US-10-425-114-44450

Query Match      2.1%  Score 7;  DB 12;  Length 83;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      272  EAFHVK 278
Db      24  EAFHVK 30

RESULT 19
US-10-424-599-165543
; Sequence 165543, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 165543
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(87)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120500C.1.pep
US-10-424-599-165543

Query Match      2.1%  Score 7;  DB 12;  Length 87;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      28  LEKQDKK 34
Db      24  LEKQDKK 30

RESULT 20
US-10-424-599-220733
; Sequence 220733, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220733
```

```
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41352C.1.pep
US-10-424-599-220733

Query Match          2.1%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLVLTLL 83
Db 67 GLVLTLL 73

RESULT 21
US-09-731-872-430
; Sequence 430, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 430
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-731-872-430

Query Match          2.1%; Score 7; DB 9; Length 99;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 KEPTTEA 65
Db 87 KEPTTEA 93

RESULT 22
US-09-876-997-430
; Sequence 430, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 430
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-876-997-430

```
Query Match          2.1%; Score 7; DB 10; Length 99;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 59 KEPTTEA 65
Db 87 KEPTTEA 93
```

```
RESULT 23
US-10-424-599-194139
; Sequence 194139, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(S3223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 194139
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17335C.1.pep
US-10-424-599-194139
```

```
Query Match          2.1%; Score 7; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 11 HKNPAEI 17
Db 95 HKNPAEI 101
```

```
RESULT 24
US-10-106-698-5213
; Sequence 5213, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides;
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5213
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

; NAME/KEY: MISC FEATURE  
; LOCATION: (13)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (49)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-5213

Query Match 2.1%; Score 7; DB 14; Length 113;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EVVSKSL 46  
Db 51 EVVSKSL 57

RESULT 25  
US-10-156-761-8084  
; Sequence 8084, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8084  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8084

Query Match 2.1%; Score 7; DB 14; Length 133;  
Best Local Similarity 100.0%; Pred. No. 11e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAAQAQE 72  
Db 82 VAAQAQE 88

RESULT 26  
US-10-282-122A-59346  
; Sequence 59346, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 59346  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-10-282-122A-59346

Query Match 2.1%; Score 7; DB 12; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKIL 21  
Db 107 AEIVKIL 113

RESULT 27  
US-10-424-599-222167  
; Sequence 222167, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 222167  
; LENGTH: 147  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_42646C.1.pcp  
US-10-424-599-222167

Query Match 2.1%; Score 7; DB 12; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DXKTDKA 38  
Db 30 DXKTDKA 36

RESULT 28  
US-10-425-114-56304

```
; Sequence 56304, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56304
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700975650_FLI.pep
US-10-425-114-56304

Query Match      2.1%; Score 7; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHEPLAK 157
DB 71 RHEPLAK 77

RESULT 29
US-10-424-599-234035
; Sequence 234035, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234035
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53358C.1.pep
US-10-424-599-234035

Query Match      2.1%; Score 7; DB 12; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 DKASEEV 42
DB 127 DKASEEV 133

RESULT 30
US-10-425-114-70029
; Sequence 70029, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
```

```
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70029
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFELB73031E03_FLI.pep
US-10-425-114-70029

Query Match      2.1%; Score 7; DB 12; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVFK 278
DB 102 EAFHVFK 108

RESULT 31
US-10-424-599-256726
; Sequence 256726, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 256726
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73847C.1.pep
US-10-424-599-256726

Query Match      2.1%; Score 7; DB 12; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLIA 84
DB 67 LLVTLIA 73

RESULT 32
US-10-424-599-173625
; Sequence 173625, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 173625
```

; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_127802C.1.pap  
; US-10-424-599-173625

Query Match 2.1%; Score 7; DB 12; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 IEFSSSF 308  
Db 7 IEFSSSF 13

## RESULT 33

US-10-424-599-221069  
; Sequence 221069, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 221069  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_41655C.1.pap  
; US-10-424-599-221069

Query Match 2.1%; Score 7; DB 12; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHEPLAK 157  
Db 88 RHEPLAK 94

## RESULT 34

US-10-425-114-38325  
; Sequence 38325, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 38325  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3242-311-C1\_FLI.pap  
; US-10-425-114-38325

Query Match 2.1%; Score 7; DB 12; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHEPLAK 157  
Db 88 RHEPLAK 94

## RESULT 35

US-10-424-599-158029  
; Sequence 158029, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 158029  
; LENGTH: 173  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_11371C.1.pap  
; US-10-424-599-158029

Query Match 2.1%; Score 7; DB 12; Length 173;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVFK 278  
Db 108 EAFHVFK 114

## RESULT 36

US-09-729-674-280  
; Sequence 280, Application US/09729674  
; Patent No. US2001003935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallie, Edward R.  
; APPLICANT: Collins-Racie, Lisa A.  
; APPLICANT: Evans, Cheryl  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Steininger II, Robert J.  
; APPLICANT: Spaulding, Vikki  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Rechtel, Kim  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: 6055-64X  
; CURRENT APPLICATION NUMBER: US/09/729,674  
; CURRENT FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: 09/539,330  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 283  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 280  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-729-674-280



Query Match 2.1%; Score 7; DB 9; Length 176;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 231 LIGELIL 237  
| | | | |  
Db 98 LLGELIL 104

## RESULT 37

US-10-108-260A-4437  
; Sequence 4437, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
; FILE REFERENCE: HL-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4437  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4437

Query Match 2.1%; Score 7; DB 15; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 27 ILEKQDK 33  
| | | | |  
Db 188 ILEKQDK 194

## RESULT 38

US-10-108-260A-4032  
; Sequence 4032, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
; FILE REFERENCE: HL-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4032  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4032

Query Match 2.1%; Score 7; DB 15; Length 199;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 73 LYSSGGL 79  
| | | | |  
Db 57 LYSSGGL 63

## RESULT 39

US-10-425-114-45029  
; Sequence 45029, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313) B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 45029  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700166135\_FLI.pep  
US-10-425-114-45029

Query Match 2.1%; Score 7; DB 12; Length 204;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 65 AVAQLAQ 71  
| | | | |  
Db 86 AVAQLAQ 92

## RESULT 40

US-09-791-279-185  
; Sequence 185, Application US/09791279  
; Publication No. US20030050456A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis  
; APPLICANT: Lind, Peter  
; TITLE OF INVENTION: No. US20030050456A1 G Protein-Coupled Receptors  
; FILE REFERENCE: 00048.US1  
; CURRENT APPLICATION NUMBER: US/09/791,279  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/184,715  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184725  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,712  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,606  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,602  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,604  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,822  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,710  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,689  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,690  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/184,716  
; PRIOR FILING DATE: 2000-02-24  
; NUMBER OF SEQ ID NOS: 220  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 185  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-279-185

Query Match 2.1%; Score 7; DB 10; Length 215;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 69 LAQELYS 75

Db 21 LAELYS 27  
|||||

## RESULT 41

US-10-094-749-2522  
; Sequence 2522, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAL, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2522  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-2522

Query Match 2.1%; Score 7; DB 15; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
|||||  
Db 21 NLAILEK 27

## RESULT 42

US-10-156-761-8653  
; Sequence 8653, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 8653  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8653

Query Match 2.1%; Score 7; DB 14; Length 227;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQLA 70  
|||||  
Db 211 EAVAQLA 217

## RESULT 43

US-10-424-599-203478  
; Sequence 203478, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David X  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 203478  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(243)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_25767C.1.pap  
US-10-424-599-203478

Query Match 2.1%; Score 7; DB 12; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EKLLQSE 220  
|||||  
Db 113 EKLLQSE 119

## RESULT 44

US-10-424-599-179206  
; Sequence 179206, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David X  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 179206  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13283C.1.pap  
US-10-424-599-179206

Query Match 2.1%; Score 7; DB 12; Length 245;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILKD 23  
| | | | |  
Db 21 IVKILKD 27

RESULT 45  
US-10-005-549-17  
; Sequence 17, Application US/10005549  
; Publication No. US20030190613A1  
; GENERAL INFORMATION:  
; APPLICANT: BOWEN, MICHAEL A.  
; APPLICANT: WU, YULI  
; APPLICANT: YANG, WEN-PIN  
; APPLICANT: FINGER, JOSHUA  
; APPLICANT: NADLER, STEVEN  
; APPLICANT: CARROLL, PAMELA  
; APPLICANT: FEDER, JOHN  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING AN ACTIVATED HUMAN  
; TITLE OF INVENTION: T-LYMPHOCYTE-DERIVED PROTEIN RELATED TO UBIQUITIN  
; TITLE OF INVENTION: CONJUGATING ENZYME  
; FILE REFERENCE: D0034np  
; CURRENT APPLICATION NUMBER: US/10/005,549  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/308,706  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/244,688  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-005-549-17

Query Match 2.1%; Score 7; DB 14; Length 245;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILEK 30  
| | | | |  
Db 138 NLAILEK 144

RESULT 46  
US-10-424-599-151467  
; Sequence 151467, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 151467  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_107798C.1.pep  
US-10-424-599-151467

Query Match 2.1%; Score 7; DB 12; Length 251;

Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 QAMKEIL 53  
| | | | |  
Db 87 QAMKEIL 93

RESULT 47  
US-10-080-170-328  
; Sequence 328, Application US/10080170  
; Publication No. US20030129601A1  
; GENERAL INFORMATION:  
; APPLICANT: COLE, S.T.  
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES  
; FILE REFERENCE: 03495.0218  
; CURRENT APPLICATION NUMBER: US/10/080,170  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: 60/270,123  
; PRIOR FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 552  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 328  
; LENGTH: 275  
; TYPE: PRT  
; ORGANISM: Mycobacterium leprae  
US-10-080-170-328

Query Match 2.1%; Score 7; DB 14; Length 275;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ILKDNLA 26  
| | | | |  
Db 74 ILKDNLA 80

RESULT 48  
US-10-282-122A-52103  
; Sequence 52103, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,948  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52103  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Clostridium acetobutylicum  
US-10-282-122A-52103

Query Match 2.1%; Score 7; DB 12; Length 282;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 RQSLKLL 232  
|||  
Db 39 RQSLKLL 45

RESULT 49  
US-10-424-599-209662  
; Sequence 209662, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 209662  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(337)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_31352C.1.pep  
US-10-424-599-209662

Query Match 2.1%; Score 7; DB 12; Length 337;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAPHVEK 278  
|||  
Db 278 EAPHVEK 284

RESULT 50  
US-10-425-114-72787  
; Sequence 72787, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 72787  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3354-066-B6\_FLI.pep  
US-10-425-114-72787

Query Match 2.1%; Score 7; DB 12; Length 338;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VSKSLQA 48  
|||  
Db 267 VSKSLQA 273

RESULT 51  
US-10-424-599-234484  
; Sequence 234484, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 234484  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_53767C.1.pep  
US-10-424-599-234484

Query Match 2.1%; Score 7; DB 12; Length 341;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LEQDKK 34  
|||  
Db 213 LEQDKK 219

RESULT 52  
US-10-282-122A-51574  
; Sequence 51574, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangau  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; SOFTWARE: PatentIn version 3.1  
; NUMBER OF SEQ ID NOS: 78614  
; SEQ ID NO 51574  
; LENGTH: 344  
; TYPE: PR1  
; ORGANISM: Clostridium acetobutylicum  
US-10-282-122A-51574

Query Match 2.1%; Score 7; DB 12; Length 344;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKILK 22  
Db 335 EIVKILK 341

RESULT 53  
US-10-425-114-46025  
; Sequence 46025, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46025  
; LENGTH: 351  
; TYPE: PR1  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700170954\_FLI.pep  
US-10-425-114-46025

Query Match 2.1%; Score 7; DB 12; Length 351;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 FDIASDA 183  
Db 191 FDIASDA 197

RESULT 54  
US-10-424-599-262896  
; Sequence 262896, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 262896  
; LENGTH: 356  
; TYPE: PR1  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_79416C.1.pep  
US-10-424-599-262896

Query Match 2.1%; Score 7; DB 12; Length 356;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234  
Db 321 SLKLGE 327

RESULT 55  
US-10-425-114-41971  
; Sequence 41971, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 41971  
; LENGTH: 360  
; TYPE: PR1  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700865271\_FLI.pep  
US-10-425-114-41971

Query Match 2.1%; Score 7; DB 12; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234  
Db 325 SLKLGE 331

RESULT 56  
US-10-425-114-47953  
; Sequence 47953, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 47953  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700909018\_FLI.pep  
US-10-425-114-47953

Query Match 2.1%; Score 7; DB 12; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLGE 234  
| | | | |  
DB 325 SLKLGE 331

RESULT 57  
US-10-369-493-42  
; Sequence 42, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 42  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: Aquifex aeolicus  
US-10-369-493-42

Query Match 2.1%; Score 7; DB 15; Length 373;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 PKLIEFL 305  
| | | | |  
DB 274 PKLIEFL 280

RESULT 58  
US-10-106-698-4458  
; Sequence 4458, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 4458  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (25)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-4458

Query Match 2.1%; Score 7; DB 14; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLALKEK 30  
| | | | |  
DB 178 NLALKEK 184

RESULT 59  
US-10-425-114-68619  
; Sequence 68619, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 68619  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17048B05\_FLI.pep  
US-10-425-114-68619

Query Match 2.1%; Score 7; DB 12; Length 391;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQEL 73  
| | | | |  
DB 375 AQLAQEL 381

RESULT 60  
US-10-425-114-37084  
; Sequence 37084, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37084
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-059-D4_FLI.pep
US-10-425-114-37084

```

Query Match 2.1%; Score 7; DB 12; Length 399;

Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHEPLAK 157

Db 319 RHEPLAK 325

RESULT 61

US-10-005-549-2

; Sequence 2, Application US/10005549

; Publication No. US20030190613A1

; GENERAL INFORMATION:

; APPLICANT: BOWEN, MICHAEL A.

; APPLICANT: WU, YULI

; APPLICANT: YANG, WEN-PIN

; APPLICANT: FINGER, JOSHUA

; APPLICANT: NADLER, STEVEN

; APPLICANT: CARROLL, PAMELA

; APPLICANT: FEDER, JOHN

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING AN ACTIVATED HUMAN

; TITLE OF INVENTION: T-LYMPHOCYTE-DERIVED PROTEIN RELATED TO UBIQUITIN

; TITLE OF INVENTION: CONJUGATING ENZYME

; FILE REFERENCE: D0034ND

; CURRENT APPLICATION NUMBER: US/10/005,549

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/308,706

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/244,688

; PRIOR FILING DATE: 2000-10-30

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 422

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-005-549-2

Query Match

2.1%; Score 7; DB 14; Length 422;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLATLEK 30

Db 226 NLATLEK 232

RESULT 62

US-10-307-928A-2

; Sequence 2, Application US/10307928A

; Publication No. US20030229016A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook, John P.

; APPLICANT: Anderson, David W.

; APPLICANT: Boldog, Ferenc L.

; APPLICANT: Burgess, Catherine E.

; APPLICANT: Catterton, Elina

; APPLICANT: Edinger, Shlomit R.

```

; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS O
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 24102-502D
; CURRENT APPLICATION NUMBER: US/10/307,928A
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/381,495
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/383,744
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/384,024
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/401,788
; PRIOR FILING DATE: 2003-08-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-928A-2

Query Match 2.1%; Score 7; DB 15; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLATLEK 30
Db 226 NLATLEK 232

RESULT 63
US-10-282-122A-59712
; Sequence 59712, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```



FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 59712  
LENGTH: 439  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-10-282-122A-59712

Query Match 2.1%; Score 7; DB 12; Length 439;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 HNFAMT 246  
|||||  
DB 117 HNFAMT 123

RESULT 64  
US-10-425-114-37583  
Sequence 37583, Application US/10425114  
Publication No. US2004003488A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 37583  
LENGTH: 448  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3960-022-A8\_FLI.pep  
US-10-425-114-37583

Query Match 2.1%; Score 7; DB 12; Length 448;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGLL 79  
|||||  
DB 95 LYSSGLL 101

RESULT 65  
US-10-238-075-859  
Sequence 859, Application US/10238075  
Publication No. US20030148324A1  
GENERAL INFORMATION:  
APPLICANT: I.N.S.E.R.M.  
TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat  
FILE REFERENCE: BLANDINE  
CURRENT APPLICATION NUMBER: US/10/238,075  
CURRENT FILING DATE: 2002-09-10  
PRIOR APPLICATION NUMBER: 0003145  
PRIOR FILING DATE: 2000-03-10  
NUMBER OF SEQ ID NOS: 1576  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 859  
LENGTH: 452  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-238-075-859

Query Match 2.1%; Score 7; DB 14; Length 452;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDAF 184  
|||||  
DB 128 DIASDAF 134

RESULT 66  
US-10-032-585-7310  
Sequence 7310, Application US/10032585  
Publication No. US20030180953A1  
GENERAL INFORMATION:  
APPLICANT: Terry, Roemer D.  
APPLICANT: Bo, Jiang  
APPLICANT: Charles, Boone  
APPLICANT: Howard, Bussey  
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
FILE REFERENCE: 10182-005-999  
CURRENT APPLICATION NUMBER: US/10/032,585  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 8000  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7310  
LENGTH: 458  
TYPE: PRT  
ORGANISM: Candida albicans  
US-10-032-585-7310

Query Match 2.1%; Score 7; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 FADEKRY 324  
|||||  
DB 332 FADEKRY 338

RESULT 67  
US-10-424-599-221067  
Sequence 221067, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: Plants and Uses Thereof for Plant Improvement

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 221067
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41653C.1.pep
US-10-424-599-221067

Query Match      2.1%; Score 7; DB 12; Length 488;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      151 RHEPLAK 157
      |||||
Db      408 RHEPLAK 414

RESULT 68
US-09-952-464A-10
; Sequence 10, Application US/09952464A
; Publication No. US20030077587A1
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Fingert, John
; TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS
; FILE REFERENCE: 21087,0017U11
; CURRENT APPLICATION NUMBER: US/09/952,464A
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/473,273
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 09/461,542
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 09/366,952
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: 09/056,285
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/822,999
; PRIOR FILING DATE: 1997-03-21
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20030077587A1e =
; OTHER INFORMATION: synthetic construct
US-09-952-464A-10

Query Match      2.1%; Score 7; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLLRDKS 266
      |||||
Db      130 NLLRDKS 136

RESULT 69
US-10-425-114-38253
; Sequence 38253, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38253
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3106-102-C2_FIL.pep
US-10-425-114-38253

Query Match      2.1%; Score 7; DB 12; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      151 RHEPLAK 157
      |||||
Db      410 RHEPLAK 416

RESULT 70
US-09-844-653-173
; Sequence 173, Application US/09844653
; Publication No. US20030054347A1
; GENERAL INFORMATION:
; APPLICANT: Richards, Julia
; APPLICANT: Rozsa, Frank
; TITLE OF INVENTION: Detecting and Treating Eye Disease
; FILE REFERENCE: UM-06105
; CURRENT APPLICATION NUMBER: US/09/844,653
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 173
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-844-653-173

Query Match      2.1%; Score 7; DB 10; Length 504;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLLRDKS 266
      |||||
Db      144 NLLRDKS 150

RESULT 71
US-09-952-464A-8
; Sequence 8, Application US/09952464A
; Publication No. US20030077587A1
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Fingert, John
; TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS
; FILE REFERENCE: 21087,0017U11
; CURRENT APPLICATION NUMBER: US/09/952,464A
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/473,273
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 09/461,542
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 09/366,952
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: 09/056,285
```

```
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/822,999
; PRIOR FILING DATE: 1997-03-21
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20030077587A1 =
US-09-952-464A-8

Query Match      2.1%; Score 7; DB 10; Length 504;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLRDKS 266
DB      144 NLRDKS 150

RESULT 72
US-10-086-135-4
; Sequence 4, Application US/10086135
; Publication No. US20020182677A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Chandrasekhar, Yasmin A.
; TITLE OF INVENTION: PANCREATIC AND OVARIAN POLYPEPTIDE
; TITLE OF INVENTION: ZSIG58
; FILE REFERENCE: 98-24
; CURRENT APPLICATION NUMBER: US/10/086,135
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/366,448
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-086-135-4

Query Match      2.1%; Score 7; DB 13; Length 504;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLRDKS 266
DB      144 NLRDKS 150

RESULT 73
US-10-244-633-32
; Sequence 32, Application US/10244633
; Publication No. US20030068640A1
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; APPLICANT: Chen, Pu
; APPLICANT: Chen, Hua
; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis,
; TITLE OF INVENTION: Prognosis And Treatment Of Glaucoma And Related
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: 07425.0057.USG1
; CURRENT APPLICATION NUMBER: US/10/244,633
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US/09/306,828
; PRIOR FILING DATE: 1999-05-07
```

```
; PRIOR APPLICATION NUMBER: US 09/227,881
; PRIOR FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 32
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-244-633-32

Query Match      2.1%; Score 7; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLRDKS 266
DB      144 NLRDKS 150

RESULT 74
US-10-425-114-55451
; Sequence 55451, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55451
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-QMFLMINSOY077B08_FLI.pep
US-10-425-114-55451

Query Match      2.1%; Score 7; DB 12; Length 508;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      151 RHEPLAK 157
DB      428 RHEPLAK 434

RESULT 75
US-09-800-729-124
; Sequence 124, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 124
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-800-729-124

Query Match 2.1%; Score 7; DB 9; Length 514;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
|||||

DB 392 VAQLAQE 398

RESULT 76

US-09-833-245-2221  
; Sequence 2221, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2221  
; LENGTH: 514  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-2221

Query Match 2.1%; Score 7; DB 11; Length 514;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
|||||

DB 392 VAQLAQE 398

RESULT 77

US-10-424-599-220386  
; Sequence 220386, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 220386  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_41039C.1.pcp  
US-10-424-599-220386

Query Match 2.1%; Score 7; DB 12; Length 524;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHEPLAK 157  
|||||

DB 425 RHEPLAK 431

RESULT 78

US-10-389-566-822  
; Sequence 822, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 822  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-389-566-822

Query Match 2.1%; Score 7; DB 16; Length 534;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKILK 22  
|||||

DB 85 EIVKILK 91

RESULT 79

US-10-389-566-1163  
; Sequence 1163, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1163  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-389-566-1163

Query Match 2.1%; Score 7; DB 16; Length 534;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EIVKILK 22  
|||||

DB 85 EIVKILK 91

RESULT 80

US-10-369-493-14077

```
; Sequence 14077, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14077
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(558)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-14077

Query Match      2.1%  Score 7;  DB 15;  Length 558;
Best Local Similarity 100.0%;  Pred. No. 4.1e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      65 AVAQLAQ 71
DB      441 AVAQLAQ 447

RESULT 81
US-10-425-114-68937
; Sequence 68937, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68937
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMOL17058D11_FLI.pap
US-10-425-114-68937

Query Match      2.1%  Score 7;  DB 12;  Length 559;
Best Local Similarity 100.0%;  Pred. No. 4.1e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      67 AQAQEL 73
DB      543 AQAQEL 549

RESULT 82
US-10-407-866-93
; Sequence 93, Application US/10407866
; Publication No. US20040002593A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: 66654-10(LJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-407-866-93

Query Match      2.1%  Score 7;  DB 15;  Length 571;
Best Local Similarity 100.0%;  Pred. No. 4.2e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      191 LTRHKVL 197
DB      394 LTRHKVL 400

RESULT 83
US-10-389-566-1261
; Sequence 1261, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1261
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-389-566-1261

Query Match      2.1%  Score 7;  DB 16;  Length 575;
Best Local Similarity 100.0%;  Pred. No. 4.2e+02;
Matches 7;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      292 EILLKQK 298
DB      34 EILLKQK 40

RESULT 84
US-10-389-566-2425
; Sequence 2425, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
```

; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2425  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-389-566-2425

Query Match 2.1%; Score 7; DB 16; Length 575;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKNQ 298  
|||  
Db 34 EILLKNQ 40

## RESULT 85

US-10-389-566-1939  
; Sequence 1939, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(53900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1939  
; LENGTH: 576  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-389-566-1939

Query Match 2.1%; Score 7; DB 16; Length 576;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKNQ 298  
|||  
Db 34 EILLKNQ 40

## RESULT 86

US-10-425-114-66875  
; Sequence 66875, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 66875  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17161E07\_FLI.pep  
US-10-425-114-66875

Query Match 2.1%; Score 7; DB 12; Length 628;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQL 73  
|||  
Db 612 AQLAQL 618

## RESULT 87

US-10-369-493-12179  
; Sequence 12179, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 12179  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Mesorhizobium loti  
US-10-369-493-12179

Query Match 2.1%; Score 7; DB 15; Length 631;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 DIASDAF 184  
|||  
Db 133 DIASDAF 139

## RESULT 88

US-10-425-114-64320  
; Sequence 64320, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 64320  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3632-018-C12\_FLI.pep

US-10-425-114-64320

Query Match 2.1%, Score 7; DB 12; Length 637;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGLL 79  
Db 284 LYSSGLL 290

RESULT 89

US-10-425-114-72475  
; Sequence 72475, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 72475  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73106C12\_FLI.pep  
US-10-425-114-72475

Query Match 2.1%, Score 7; DB 12; Length 637;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGLL 79  
Db 284 LYSSGLL 290

RESULT 90

US-10-086-464-8  
; Sequence 8, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 731  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-086-464-8

Query Match 2.1%, Score 7; DB 13; Length 731;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 195 KVLVADF 201  
Db 519 KVLVADF 525

RESULT 91

US-10-369-493-2071  
; Sequence 2071, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2071  
; LENGTH: 762  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2071

Query Match 2.1%, Score 7; DB 15; Length 762;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 ADQLID 90  
Db 321 ADQLID 327

RESULT 92

US-10-312-088-34  
; Sequence 34, Application US/10312088  
; Publication No. US20030219862A1  
; GENERAL INFORMATION:  
; APPLICANT: Agarwal, pankaj  
; APPLICANT: Cogswell, John P.  
; APPLICANT: Kabnic, Karen S.  
; APPLICANT: Lai, Ying-Ta  
; APPLICANT: Martensen, Shelby A.  
; APPLICANT: Murdock, Paul R.  
; APPLICANT: Smith, Randall F.  
; APPLICANT: Strum, Jay C.  
; APPLICANT: Xiang, Zhaoying  
; APPLICANT: Xie, Qing  
; APPLICANT: Rizni, Safia K.  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP50029  
; CURRENT APPLICATION NUMBER: US/10/312,088  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: PCT/US01/19929  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213,161  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/213,156  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 855  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-312-088-34





```
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69986
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69986

Query Match          2.1%; Score 7; DB 12; Length 891;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AQAQEL 73
Db 639 AQAQEL 645
|||||

RESULT 96
US-10-334-143-27
; Sequence 27, Application US/10334143
; Publication No. US2004009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; FILE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-27

Query Match          2.1%; Score 7; DB 15; Length 893;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 EEVSKSL 46
Db 831 EEVSKSL 837
|||||

RESULT 97
US-10-282-122A-69548
; Sequence 69548, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69548
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69548

Query Match          2.1%; Score 7; DB 12; Length 1053;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 TLIALDLQ 87
Db 647 TLIALDLQ 653
|||||

RESULT 98
US-10-369-493-3006
; Sequence 3006, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3006
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-3006

Query Match          2.1%; Score 7; DB 15; Length 1170;
```

Best Local Similarity 100.0%; Pred. No. 7.9e+02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 170 KYVELST 176  
Db 310 KYVELST 316

## RESULT 99

US-09-800-729-89  
; Sequence 89, Application US/09800729  
; Patent No. US20020068319A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: PZ044PI  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 89  
; LENGTH: 1745  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-89

Query Match 2.1%; Score 7; DB 9; Length 1745;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 66 VAQLAQE 72  
Db 1084 VAQLAQE 1090

## RESULT 100

US-09-833-245-2222  
; Sequence 2222, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2222  
; LENGTH: 1745  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-2222

Query Match 2.1%; Score 7; DB 11; Length 1745;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 66 VAQLAQE 72  
Db 1084 VAQLAQE 1090

## RESULT 101

## US-10-044-807-2

; Sequence 2, Application US/10044807  
; Publication No. US20020165187A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. US20020165187A1 Human Protease and Polynucleotides Encodir  
; FILE REFERENCE: LEX-0298-USA  
; CURRENT APPLICATION NUMBER: US/10/044,807  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: US 60/261,684  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1762  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-044-807-2

Query Match 2.1%; Score 7; DB 13; Length 1762;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 66 VAQLAQE 72  
Db 1101 VAQLAQE 1107

## RESULT 102

US-09-764-853-554  
; Sequence 554, Application US/09764853  
; Patent No. US20020090672A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJZ06  
; CURRENT APPLICATION NUMBER: US/09/764,853  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 939  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 554  
; LENGTH: 1766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (533)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-853-554

Query Match 2.1%; Score 7; DB 9; Length 1766;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 66 VAQLAQE 72  
Db 1105 VAQLAQE 1111

## RESULT 103

US-10-399-385-5  
; Sequence 5, Application US/10399385  
; Publication No. US20040033513A1  
; GENERAL INFORMATION:  
; APPLICANT: Camaschella, Clara  
; APPLICANT: Kury, Friedrich  
; APPLICANT: Oberkannis, Christian  
; TITLE OF INVENTION: METHOD AND PROBE FOR GENETIC DIAGNOSIS OF HEREDITARY HAEMOCHROMA  
; FILE REFERENCE: 37396/VIENNALAB  
; CURRENT APPLICATION NUMBER: US/10/399,385  
; CURRENT FILING DATE: 2003-04-16

; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-385-5

Query Match 1.8%; Score 6; DB 12; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 AVAQLA 70  
Db 2 AVAQLA 7

## RESULT 104

US-10-261-798-51  
; Sequence 51, Application US/10261798  
; Publication No. US20030144477A1  
; GENERAL INFORMATION:  
; APPLICANT: Spack, Edward  
; APPLICANT: Arimilli, Subhashini  
; APPLICANT: Deshpande, Shrikant  
; APPLICANT: Wehner, Nancy  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Immunodominant Acetylcholine Receptor Alpha  
; TITLE OF INVENTION: Subunit Peptide:MHC Complexes  
; FILE REFERENCE: 014058-015810US  
; CURRENT APPLICATION NUMBER: US/10/261,798  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: US 60/327,495  
; PRIOR FILING DATE: 2001-10-04  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 51  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: acetylcholine receptor (Achr) alpha peptide  
; OTHER INFORMATION: 330-343  
US-10-261-798-51

Query Match 1.8%; Score 6; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EKODKK 34  
Db 6 EKODKK 11

## RESULT 105

US-10-399-385-4  
; Sequence 4, Application US/10399385  
; Publication No. US20040033513A1  
; GENERAL INFORMATION:  
; APPLICANT: Camaschella, Clara  
; APPLICANT: Kury, Friedrich  
; APPLICANT: Oberkannig, Christian  
; TITLE OF INVENTION: METHOD AND PROBE FOR GENETIC DIAGNOSIS OF HEREDITARY HAEMOCHROMATOSIS  
; FILE REFERENCE: 37396/VIENNALAB  
; CURRENT APPLICATION NUMBER: US/10/399,385  
; CURRENT FILING DATE: 2003-04-16  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-399-385-4

Query Match 1.8%; Score 6; DB 12; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 AVAQLA 70  
Db 6 AVAQLA 11

## RESULT 106

US-09-864-761-46460  
; Sequence 46460, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wenheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 46460  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO Z83840.7  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.3  
US-09-864-761-46460

Query Match 1.8%; Score 6; DB 9; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 SKPENL 255  
| | | | |  
DB 11 SKPENL 16

## RESULT 107

US-09-774-639-312  
; Sequence 312, Application US/09774639  
; Publication No. US2003000355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P1  
; CURRENT APPLICATION NUMBER: US/09/774,639  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 371  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 312  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-774-639-312

Query Match 1.8%; Score 6; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KKTDKA 38  
| | | | |  
DB 8 KKTDKA 13

## RESULT 108

US-09-969-730-239  
; Sequence 239, Application US/09969730  
; Publication No. US2003005443A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P2  
; CURRENT APPLICATION NUMBER: US/09/969,730  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,732  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,366  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,364  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,367  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,365

; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,731  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,557  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,563  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/055,970  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,986  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,311  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,808  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,803  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,809  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,806  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/055,310  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,798  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/055,309  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/055,312  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/054,807  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: 60/055,386  
; PRIOR FILING DATE: 1997-08-05  
; NUMBER OF SEQ ID NOS: 373  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 239  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-969-730-239

Query Match 1.8%; Score 6; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KKTDKA 38  
| | | | |  
DB 8 KKTDKA 13

## RESULT 109

US-10-621-363-239  
; Sequence 239, Application US/10621363  
; Publication No. US20040023283A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P2C1  
; CURRENT APPLICATION NUMBER: US/10/621,363  
; CURRENT FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 09/969,730  
; PRIOR FILING DATE: 2001-10-06  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45850
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008998.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: AA378559.1, EVALUAE 4.00e-13
US-09-864-761-45850

Query Match 1.8%; Score 6; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels

Qy 287 TQPIVE 292
Db 1 TQPIVE 6
|||||

RESULT 111
US-10-424-599-230826
; Sequence 230826, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 230826
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_50457C.1.pap
US-10-424-599-230826

Query Match 1.8%; Score 6; DB 12; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels

Qy 76 SGLLVT 81
Db 23 SGLLVT 28
|||||

RESULT 112
US-09-917-340-62
; Sequence 62, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340

```

```

; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-917-340-62

```

```

Query Match      1.8%; Score 6; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      106 RROIGT 111
      |||||
Db      25 RROIGT 30

```

```

RESULT 113
US-09-917-340-63
; Sequence 63, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TELANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-917-340-63

```

```

Query Match      1.8%; Score 6; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      106 RROIGT 111
      |||||
Db      25 RROIGT 30

```

```

RESULT 114
US-09-864-761-46961
; Sequence 46961, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

```

```

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46961
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004832.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P49193, EVALUATE 1.90e-01
US-09-864-761-46961

```

```

Query Match      1.8%; Score 6; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      36 DKASEE 41
      |||||
Db      26 DKASEE 31

```

```

RESULT 115
US-09-864-761-33835
; Sequence 33835, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

```



PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33835  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC007846.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
NAME/KEY: unsure  
LOCATION: 20  
NAME/KEY: unsure  
LOCATION: 30  
US-09-864-761-33835

Query Match 1.8%; Score 6; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
Db 35 GLLVTL 40

## RESULT 116

US-09-764-891-3628  
Sequence 3628, Application US/09764891  
Publication No. US20030077808A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3628  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (9)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-3628

Query Match 1.8%; Score 6; DB 10; Length 43;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 FLSSEQ 309  
Db 27 FLSSEQ 32

## RESULT 117

US-10-291-265-372  
Sequence 372, Application US/10291265  
Publication No. US20030232054A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: No. US20030232054A1 Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-017 (785)  
CURRENT APPLICATION NUMBER: US/10/291,265  
CURRENT FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 09/491,404  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 09/617,746  
PRIOR FILING DATE: 2000-07-17  
PRIOR APPLICATION NUMBER: 09/631,451  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 09/633,870  
PRIOR FILING DATE: 2000-09-15  
NUMBER OF SEQ ID NOS: 944  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 372  
LENGTH: 44  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-291-265-372

Query Match 1.8%; Score 6; DB 15; Length 44;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 KIILFS 162  
Db 36 KIILFS 41

## RESULT 118

US-10-424-599-157183  
Sequence 157183, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei

;/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
;/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;/ FILE REFERENCE: 38-21(53223)B  
;/ CURRENT APPLICATION NUMBER: US/10/424,599  
;/ CURRENT FILING DATE: 2003-04-28  
;/ NUMBER OF SEQ ID NOS: 285684  
;/ SEQ ID NO 157183  
;/ LENGTH: 45  
;/ TYPE: PRT  
;/ ORGANISM: Glycine max  
;/ FEATURE:  
;/ OTHER INFORMATION: Clone ID: PAT\_MRT3847\_112958C.1.pep  
US-10-424-599-157183

Query Match 1.8%; Score 6; DB 12; Length 45;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 LFMLLK 131  
Db 6 LFMLLK 11

RESULT 119  
US-10-424-599-159316  
;/ Sequence 159316, Application US/10424599  
;/ Publication No. US20040031072A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: La Rosa Thomas J  
;/ APPLICANT: Kovalic David K  
;/ APPLICANT: Zhou Yihua  
;/ APPLICANT: Cao Yongwei  
;/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
;/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;/ FILE REFERENCE: 38-21(53223)B  
;/ CURRENT APPLICATION NUMBER: US/10/424,599  
;/ CURRENT FILING DATE: 2003-04-28  
;/ NUMBER OF SEQ ID NOS: 285684  
;/ SEQ ID NO 159316  
;/ LENGTH: 46  
;/ TYPE: PRT  
;/ ORGANISM: Glycine max  
;/ FEATURE:  
;/ NAME/KEY: unsure  
;/ LOCATION: (1)..(46)  
;/ OTHER INFORMATION: unsure at all Xaa locations  
;/ FEATURE:  
;/ OTHER INFORMATION: Clone ID: PAT\_MRT3847\_114883C.1.pep  
US-10-424-599-159316

Query Match 1.8%; Score 6; DB 12; Length 46;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RRQIGT 111  
Db 26 RRQIGT 31

RESULT 120  
US-09-764-872-445  
;/ Sequence 445, Application US/09764872  
;/ Publication No. US20030050231A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Rosen et al.  
;/ APPLICANT: Cao Yongwei  
;/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
;/ FILE REFERENCE: PA125  
;/ CURRENT APPLICATION NUMBER: US/09/764,872  
;/ CURRENT FILING DATE: 2001-01-17  
;/ Prior application data removed - consult PALM or file wrapper  
;/ NUMBER OF SEQ ID NOS: 957  
;/ SOFTWARE: PatentIn Ver. 2.0  
;/ SEQ ID NO 445

;/ LENGTH: 47  
;/ TYPE: PRT  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: SITE  
;/ LOCATION: (43)  
;/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-872-445

Query Match 1.8%; Score 6; DB 10; Length 47;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107  
Db 20 NNILRR 25

RESULT 121  
US-10-424-599-188181  
;/ Sequence 188181, Application US/10424599  
;/ Publication No. US20040031072A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: La Rosa Thomas J  
;/ APPLICANT: Kovalic David K  
;/ APPLICANT: Zhou Yihua  
;/ APPLICANT: Cao Yongwei  
;/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
;/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;/ FILE REFERENCE: 38-21(53223)B  
;/ CURRENT APPLICATION NUMBER: US/10/424,599  
;/ CURRENT FILING DATE: 2003-04-28  
;/ NUMBER OF SEQ ID NOS: 285684  
;/ SEQ ID NO 188181  
;/ LENGTH: 48  
;/ TYPE: PRT  
;/ ORGANISM: Glycine max  
;/ FEATURE:  
;/ OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140941C.1.pep  
US-10-424-599-188181

Query Match 1.8%; Score 6; DB 12; Length 48;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 MLKGY 133  
Db 30 MLKGY 35

RESULT 122  
US-10-424-599-269667  
;/ Sequence 269667, Application US/10424599  
;/ Publication No. US20040031072A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: La Rosa Thomas J  
;/ APPLICANT: Kovalic David K  
;/ APPLICANT: Zhou Yihua  
;/ APPLICANT: Cao Yongwei  
;/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
;/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;/ FILE REFERENCE: 38-21(53223)B  
;/ CURRENT APPLICATION NUMBER: US/10/424,599  
;/ CURRENT FILING DATE: 2003-04-28  
;/ NUMBER OF SEQ ID NOS: 285684  
;/ SEQ ID NO 269667  
;/ LENGTH: 48  
;/ TYPE: PRT  
;/ ORGANISM: Glycine max  
;/ FEATURE:  
;/ OTHER INFORMATION: Clone ID: PAT\_MRT3847\_85529C.1.pep  
US-10-424-599-269667

Query Match 1.8%; Score 6; DB 12; Length 48;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 VTLIAD 85  
Db 33 VTLIAD 38

RESULT 123  
US-10-424-599-174437  
; Sequence 174437, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 174437  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_128535C.1.pep  
US-10-424-599-174437

Query Match 1.8%; Score 6; DB 12; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 LKKTAP 337  
Db 28 LKKTAP 33

RESULT 124  
US-10-424-599-186215  
; Sequence 186215, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 186215  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_139165C.1.pep  
US-10-424-599-186215

Query Match 1.8%; Score 6; DB 12; Length 49;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 KYVELS 175  
Db 3 KYVELS 8

RESULT 125  
US-10-106-698-6116  
; Sequence 6116, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005F1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6116  
; LENGTH: 50  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-6116

Query Match 1.8%; Score 6; DB 14; Length 50;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEAVAQ 68  
Db 15 TEAVAQ 20

RESULT 126  
US-10-424-599-223354  
; Sequence 223354, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 223354  
; LENGTH: 51  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_4371C.1.pep  
US-10-424-599-223354

Query Match 1.8%; Score 6; DB 12; Length 51;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295  
Db 24 IVEILL 29

RESULT 127  
US-10-424-599-249890  
; Sequence 249890, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 249890  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_6767C.1.pep  
US-10-424-599-249890

Query Match 1.8%; Score 6; DB 12; Length 52;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 297 NOPKLI 302  
Db 36 NOPKLI 41

RESULT 128  
US-10-156-761-15079  
; Sequence 15079, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 15079  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-15079

Query Match 1.8%; Score 6; DB 14; Length 52;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 LYSSGL 78  
Db 9 LYSSGL 14

RESULT 129  
US-10-424-599-192699  
; Sequence 192699, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 192699  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_16029C.1.pep  
US-10-424-599-192699

Query Match 1.8%; Score 6; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LAILEK 30  
Db 30 LAILEK 35

RESULT 130  
US-10-424-599-247013  
; Sequence 247013, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 247013  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)...(53)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_65082C.1.pep  
US-10-424-599-247013

Query Match 1.8%; Score 6; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 ADFLEQ 204  
Db 43 ADFLEQ 48

RESULT 131  
US-10-424-599-149472  
; Sequence 149472, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 149472  
; LENGTH: 54  
; TYPE: PRT

```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105997C.1.pep
US-10-424-599-149472

Query Match
  1.8% Score 6; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 LFWLLK 131
Db 46 LFWLLK 51

RESULT 132
US-10-424-599-182688
; Sequence 182688, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 182688
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135980C.1.pep
US-10-424-599-182688

Query Match
  1.8% Score 6; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 LKKTAP 337
Db 39 LKKTAP 44

RESULT 133
US-09-908-711-83
; Sequence 83, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,905
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01354
```

```
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,891
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01339
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01340
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,874
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01334
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,898
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01320
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,853
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01349
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,902
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01239
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01348
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,882
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01347
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,896
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01307
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01341
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,856
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01336
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,868
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01312
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-908-711-83
```

Query Match 1.8% Score 6; DB 9; Length 56;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
|||||

Db

6 VKILKD 11

## RESULT 134

US-09-864-761-49022  
; Sequence 49022, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US 09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 49022  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC007358.2  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1  
; OTHER INFORMATION: EST HUMAN HIT: A1276750.1, EVALUATE 1.00e-25  
; OTHER INFORMATION: SWISSPROT HIT: P54660, EVALUATE 7.70e+00

## US-09-864-761-49022

Query Match 1.8%; Score 6; DB 9; Length 56;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 EFLSSF 308  
Db 19 EFLSSF 24

## RESULT 135

US-09-764-891-3081  
; Sequence 3081, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3081  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (12)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

## US-09-764-891-3081

Query Match 1.8%; Score 6; DB 10; Length 56;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

18 VKILKD 23

## Db

6 VKILKD 11

## RESULT 136

US-10-424-599-267396  
; Sequence 267396, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 267396  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_B3480C.1.pcp  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 49022

## US-10-424-599-267396

Query Match 1.8%; Score 6; DB 12; Length 56;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

246 TKYISK 251

## Db

44 TKYISK 49

## RESULT 137

US-10-424-599-178427  
; Sequence 178427, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 178427  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_132137C.1.pep  
US-10-424-599-178427

Query Match 1.8%; Score 6; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 RHKVLV 198  
Db 20 RHKVLV 25

RESULT 138  
US-10-029-386-27862  
; Sequence 27862, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: ABOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 27862  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL137007.4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.98  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.98  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.62  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
; OTHER INFORMATION: SWISSPROT HIT: Q9NZ52, EVALUATE 2.10e+00  
US-10-029-386-27862

Query Match 1.8%; Score 6; DB 14; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLLGE 234  
Db 24 LKLLGE 29

RESULT 139  
US-10-127-641-5  
; Sequence 5, Application US/10127641  
; Publication No. US20030099931A1  
; GENERAL INFORMATION:  
; APPLICANT: Hart, Mary Kate  
; TITLE OF INVENTION: Prophylactic and Therapeutic Monoclonal Antibodies  
; FILE REFERENCE: 003/253/SAP  
; CURRENT APPLICATION NUMBER: US/10/127,641

; CURRENT FILING DATE: 2002-09-10  
; PRIOR APPLICATION NUMBER: US 60/285,601  
; PRIOR FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Apple Macintosh Microsoft Word 6.0  
; SEQ ID NO 5  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Eastern equine encephalitis virus  
US-10-127-641-5

Query Match 1.8%; Score 6; DB 14; Length 58;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 LEQNYD 207  
Db 37 LEQNYD 42

RESULT 140  
US-10-424-599-195329  
; Sequence 195329, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 195329  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_18409C.1.pep  
US-10-424-599-195329

Query Match 1.8%; Score 6; DB 12; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PLFSKS 10  
Db 7 PLFSKS 12

RESULT 141  
US-10-424-599-259644  
; Sequence 259644, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 259644  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_76484C.1.pep

US-10-424-599-259644

Query Match 1.8%; Score 6; DB 12; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 KILF5 162  
DB 31 KILF5 36

RESULT 142

US-10-424-599-191627  
; Sequence 191627, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 191627  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_1505C.1.pep  
US-10-424-599-191627

Query Match 1.8%; Score 6; DB 12; Length 62;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKILF 161  
DB 48 AKILF 53

RESULT 143

US-10-424-599-211257  
; Sequence 211257, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 211257  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_32792C.1.pep  
US-10-424-599-211257

Query Match 1.8%; Score 6; DB 12; Length 62;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 SKSLQ 48  
DB 18 SKSLQ 23

RESULT 144

US-10-424-599-209297  
; Sequence 209297, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 209297  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_31022C.1.pep  
US-10-424-599-209297

Query Match 1.8%; Score 6; DB 12; Length 63;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 PIVEIL 294  
DB 33 PIVEIL 38

RESULT 145

US-10-424-599-285560  
; Sequence 285560, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 285560  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_99890C.1.pep  
US-10-424-599-285560

Query Match 1.8%; Score 6; DB 12; Length 63;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 QLIDFE 92  
DB 8 QLIDFE 13

RESULT 146

US-09-864-761-47397  
; Sequence 47397, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.



APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 47397  
LENGTH: 64  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ACO19083.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.56  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.54  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.53  
OTHER INFORMATION: SWISSPROT HIT: P29775, EVALUE 2.00e-13  
OTHER INFORMATION: EST\_HUMAN HIT: A1991109.1, EVALUE 2.00e-32  
US-09-864-761-47397

Query Match 1.8%; Score 6; DB 9; Length 64;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69  
DB 40 EAVAQL 45

RESULT 147  
US-10-424-599-159823  
Sequence 159823, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 159823  
LENGTH: 64  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_115339C.1.pap  
US-10-424-599-159823

Query Match 1.8%; Score 6; DB 12; Length 64;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 QSLKLL 232  
DB 19 QSLKLL 24

RESULT 148  
US-10-335-977-5580  
Sequence 5580, Application US/10335977  
Publication No. US20040052799A1  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS SMITH et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 10031  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/335,977  
FILING DATE: 30-Dec-2002  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/993,002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 5580:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...64

SEQUENCE DESCRIPTION: SEQ ID NO: 5580:  
US-10-335-977-5580

Query Match 1.8%; Score 6; DB 12; Length 64;  
Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 VKILKD 23  
Db 35 VKILKD 40

## RESULT 149

US-10-424-599-215660  
; Sequence 215660, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 215660  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_3676C.1.pep  
US-10-424-599-215660

Query Match 1.8%; Score 6; DB 12; Length 65;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 AKILF 161  
Db 22 AKILF 27

## RESULT 150

US-10-424-599-165707  
; Sequence 165707, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 165707  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(66)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_120649C.1.pep  
US-10-424-599-165707

Query Match 1.8%; Score 6; DB 12; Length 66;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 VKILKD 23  
Db 14 VKILKD 19

## RESULT 151

US-10-424-599-191683  
; Sequence 191683, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 191683  
; LENGTH: 67  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_1510C.1.pep  
US-10-424-599-191683

Query Match 1.8%; Score 6; DB 12; Length 67;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 229 LKLLGE 234  
Db 15 LKLLGE 20

## RESULT 152

US-10-424-599-211264  
; Sequence 211264, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 211264  
; LENGTH: 67  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(67)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_32799C.1.pep  
US-10-424-599-211264

Query Match 1.8%; Score 6; DB 12; Length 67;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 IFNNIL 105  
Db 8 IFNNIL 13

```
RESULT 153
US-10-424-599-144751
; Sequence 144751, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144751
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101729C.1.pep
US-10-424-599-144751

Query Match      1.8%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295
Db 48 IVEILL 53

RESULT 154
US-10-424-599-195946
; Sequence 195946, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 195946
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_18966C.1.pep
US-10-424-599-195946

Query Match      1.8%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 KDLLTR 193
Db 34 KDLLTR 39

RESULT 155
US-10-424-599-202460
; Sequence 202460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 202460
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_24847C.1.pep
US-10-424-599-202460

Query Match      1.8%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 IEFLSS 307
Db 62 IEFLSS 67

RESULT 156
US-10-424-599-262460
; Sequence 262460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262460
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_79023C.1.pep
US-10-424-599-262460

Query Match      1.8%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 KSPNIQ 270
Db 1 KSPNIQ 6

RESULT 157
US-10-221-278-242
; Sequence 242, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
```

; PRIOR APPLICATION NUMBER: 09/596,193  
; PRIOR FILING DATE: 2000-06-17  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; NUMBER OF SEQ ID NOS: 752  
; SEQ ID NO 242  
; LENGTH: 71  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-221-278-242

Query Match 1.8%; Score 6; DB 12; Length 71;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
|||||  
Db 24 SLKLLG 29

## RESULT 158

US-10-291-172-242  
; Sequence 242, Application US/10291172  
; Publication No. US20030228584A1  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-045  
; CURRENT APPLICATION NUMBER: US/10/291,172  
; CURRENT FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/693,267  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/665,363  
; PRIOR FILING DATE: 2000-09-19  
; PRIOR APPLICATION NUMBER: 09/616,847  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 09/596,193  
; PRIOR FILING DATE: 2000-06-17  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; NUMBER OF SEQ ID NOS: 752  
; SEQ ID NO 242  
; LENGTH: 71  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-172-242

Query Match 1.8%; Score 6; DB 15; Length 71;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
|||||  
Db 24 SLKLLG 29

## RESULT 159

US-10-282-122A-76423  
; Sequence 76423, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangau  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 76423  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Treponema pallidum  
US-10-282-122A-76423

Query Match 1.8%; Score 6; DB 12; Length 72;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 ILRRQI 109  
|||||  
Db 45 ILRRQI 50

## RESULT 160

US-09-864-408A-998  
; Sequence 998, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shmukets, Richard A.

; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enc  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 998  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-864-408A-998

Query Match 1.8%; Score 6; DB 11; Length 73;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 FLSSEF 309  
|||||

```
Db      20 FLSSFQ 25

RESULT 161
US-10-424-599-203219
; Sequence 203219, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 203219
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(73)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25531C.1.pep
US-10-424-599-203219

Query Match      1.8%; Score 6; DB 12; Length 73;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      65 AVAQLA 70
      |||||
Db      13 AVAQLA 18

RESULT 162
US-10-424-599-244014
; Sequence 244014, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 244014
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_62372C.1.pep
US-10-424-599-244014

Query Match      1.8%; Score 6; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36 DKASEE 41
      |||||
Db      8 DKASEE 13

RESULT 163
US-10-425-114-37263
; Sequence 37263, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37263
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB43-49-A12_FLI.pep
US-10-425-114-37263

Query Match      1.8%; Score 6; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      293 ILLKNQ 298
      |||||
Db      64 ILLKNQ 69

RESULT 164
US-10-425-114-56749
; Sequence 56749, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56749
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017036C08_FLI.pep
US-10-425-114-56749

Query Match      1.8%; Score 6; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      277 FKVFVA 282
      |||||
Db      14 FKVFVA 19

RESULT 165
US-09-864-408A-2574
; Sequence 2574, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
```

```
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enco
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2574
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-2574

Query Match      1.8%; Score 6; DB 11; Length 76;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 NLAILE 29
Db      16 NLAILE 21

RESULT 166
US-09-867-550-574
; Sequence 574, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US98 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-574

Query Match      1.8%; Score 6; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      65 AVAQILA 70
Db      3 AVAQILA 8

RESULT 167
US-10-424-599-253313
; Sequence 253313, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 253313
```

```
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_70767C.1.pep
US-10-424-599-253313

Query Match      1.8%; Score 6; DB 12; Length 78;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      158 IILFSN 163
Db      70 IILFSN 75

RESULT 168
US-10-002-344A-200
; Sequence 200, Application US/10002344A
; Publication No. US20020172959A1
; GENERAL INFORMATION:
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenchua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
; FILE REFERENCE: DEX-0241
; CURRENT APPLICATION NUMBER: US/10/002,344A
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/242,998
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 200
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-002-344A-200

Query Match      1.8%; Score 6; DB 13; Length 78;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 LFSKSH 11
Db      29 LFSKSH 34

RESULT 169
US-10-424-599-162228
; Sequence 162228, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162228
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117509C.1.pep
US-10-424-599-162228

Query Match      1.8%; Score 6; DB 12; Length 80;
```

Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LKDNLA 26  
Db 9 LKDNLA 14

RESULT 170  
US-10-424-599-198296  
; Sequence 198296, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 198296  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_21086C.1.pep  
US-10-424-599-198296

Query Match 1.8%; Score 6; DB 12; Length 80;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 LIEFLS 306  
Db 16 LIEFLS 21

RESULT 171  
US-10-424-599-244639  
; Sequence 244639, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 244639  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_62936C.1.pep  
US-10-424-599-244639

Query Match 1.8%; Score 6; DB 12; Length 80;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 ROIGTR 112  
Db 60 ROIGTR 65

RESULT 172

US-09-738-626-4641  
; Sequence 4641, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIALI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4641  
; LENGTH: 81  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4641

Query Match 1.8%; Score 6; DB 9; Length 81;  
Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EIVKIL 21  
Db 33 EIVKIL 38

RESULT 173  
US-09-858-664A-32  
; Sequence 32, Application US/09858664A  
; Patent No. US20020072491A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui, et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000927-CIP  
; CURRENT APPLICATION NUMBER: US/09/858,664A  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/711,134  
; PRIOR FILING DATE: 2000-11-11  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-858-664A-32

Query Match 1.8%; Score 6; DB 9; Length 82;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 292 EILLKN 297  
Db 64 EILLKN 69

RESULT 174





US-09-764-891-3412

Query Match 1.8%; Score 6; DB 10; Length 83;  
Best Local Similarity 100.0%; Pred.No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 SSGLLV 80  
Db 21 SSGLLV 26

RESULT 178

US-10-424-599-231338  
; Sequence 231338, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 231338  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_5091C.1.pbp  
US-10-424-599-231338

Query Match 1.8%; Score 6; DB 12; Length 83;  
Best Local Similarity 100.0%; Pred.No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 SSFQKE 311  
Db 39 SSFQKE 44

RESULT 179

US-09-864-761-37566  
; Sequence 37566, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37566  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC011456.2  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5  
; OTHER INFORMATION: EST HUMAN HIT: BF698187.1, EVALUATE 9.30e-01  
; OTHER INFORMATION: SWISSPROT HIT: P13615, EVALUATE 3.50e-01  
US-09-864-761-37566

Query Match 1.8%; Score 6; DB 9; Length 84;  
Best Local Similarity 100.0%; Pred.No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 VELSTP 177  
Db 76 VELSTP 81

RESULT 180

US-10-424-599-245645  
; Sequence 245645, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 245645  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_6384C.1.pbp  
US-10-424-599-245645

Query Match 1.8%; Score 6; DB 12; Length 84;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 SLKLLG 233  
|||  
DB 34 SLKLLG 39

## RESULT 181

US-10-424-599-186473  
; Sequence 186473, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 186473  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(85)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_139399C.1.pep  
US-10-424-599-186473

Query Match 1.8%; Score 6; DB 12; Length 85;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 EIVKIL 21  
|||  
DB 52 EIVKIL 57

## RESULT 182

US-10-424-599-211413  
; Sequence 211413, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 211413  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_32932C.1.pep  
US-10-424-599-211413

Query Match 1.8%; Score 6; DB 12; Length 85;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 KMPLFS 8  
|||  
DB 12 KMPLFS 17

## RESULT 183

US-09-925-301-1603  
; Sequence 1603, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1603  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (30)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (45)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (62)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (63)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (73)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (80)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (81)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-301-1603

Query Match 1.8%; Score 6; DB 9; Length 86;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LFSKSH 11  
|||  
DB 13 LFSKSH 18

## RESULT 184

US-09-864-408A-8152  
; Sequence 8152, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shimkets, Richard A.  
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enc  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 8152  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-864-408A-8152

Query Match 1.8%; Score 6; DB 11; Length 87;  
Best Local Similarity 100.0%; Pred. No. 8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 TEVAQA 68  
| | | | |  
Db 63 TEVAQA 68

## RESULT 185

US-10-424-599-153190  
; Sequence 153190, Application US/10424599  
; Publication No. US20040031072A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 153190

; LENGTH: 87

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_109355C.1.pep

US-10-424-599-153190

Query Match 1.8%; Score 6; DB 12; Length 87;  
Best Local Similarity 100.0%; Pred. No. 8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 KEPPE 64  
| | | | |  
Db 15 KEPPE 20

## RESULT 186

US-09-764-887-257

; Sequence 257, Application US/09764887

; Patent No. US20020042096A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PALL3

; CURRENT APPLICATION NUMBER: US/09/764,887

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 658

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 257

; LENGTH: 88

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (87)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-887-257

Query Match 1.8%; Score 6; DB 9; Length 88;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 VKILKD 23  
| | | | |  
Db 75 VKILKD 80

## RESULT 187

US-10-424-599-165968

; Sequence 165968, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 165968

; LENGTH: 88

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_120885C.1.pep

US-10-424-599-165968

Query Match 1.8%; Score 6; DB 12; Length 88;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 VEILLK 296  
| | | | |  
Db 22 VEILLK 27

## RESULT 188

US-10-424-599-251249

; Sequence 251249, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 251249

; LENGTH: 88

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_68906C.1.pep

US-10-424-599-251249

Query Match 1.8%; Score 6; DB 12; Length 88;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 QSLKLL 232  
| | | | |  
Db 49 QSLKLL 54

## RESULT 189

US-10-335-977-8935

; Sequence 8935, Application US/10335977

```
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 8935:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...88
; SEQUENCE DESCRIPTION: SEQ ID NO: 8935:
US-10-335-977-8935
Query Match 1.8%; Score 6; DB 12; Length 88;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 329 IRDLKK 334
Db 35 IRDLKK 40
RESULT 190
US-10-073-961-257
; Sequence 257, Application US/10073961
; Publication No. US2003007602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
```

; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264

; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/233,063  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,397  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,399  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,401  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/241,808  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,826  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,786  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,221  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,475  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/231,243  
; PRIOR FILING DATE: 2000-09-08

Query Match 1.8%; Score 6; DB 14; Length 88;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 VKILKD 23  
| | | | |  
Db 75 VKILKD 80

RESULT 191  
US-10-424-599-268688  
; Sequence 268688, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 268688  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_84648C.1.pep  
US-10-424-599-268688

Query Match 1.8%; Score 6; DB 12; Length 89;

Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 TRSPTV 116  
Db 28 TRSPTV 33

RESULT 192  
US-10-335-977-8936  
; Sequence 8936, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/335,977  
; FILING DATE: 30-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,002  
; FILING DATE: 17-DEC-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragoras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: GTN-018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 8936:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 89 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...89  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8936:  
US-10-335-977-8936

Query Match 1.8%; Score 6; DB 12; Length 89;  
Best Local Similarity 100.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 IRDLKK 334  
Db 36 IRDLKK 41

RESULT 193  
US-09-882-227-114  
; Sequence 114, Application US/09882227  
; Publication No. US20030158396A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold

; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean-Francois  
; APPLICANT: Coomen, Raymond P.  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in the  
; FILE REFERENCE: 06132/047002  
; CURRENT APPLICATION NUMBER: US/09/882,227  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 08/902,615  
; PRIOR FILING DATE: 1997-07-29  
; NUMBER OF SEQ ID NOS: 638  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori  
; US-09-882-227-114

Query Match 1.8%; Score 6; DB 10; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 IRDLKK 334  
Db 37 IRDLKK 42

RESULT 194  
US-10-424-599-163560  
; Sequence 163560, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 163560  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT MRT3847\_118713C.1.pep  
US-10-424-599-163560

Query Match 1.8%; Score 6; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 IASDAF 184  
Db 39 IASDAF 44

RESULT 195  
US-10-424-599-186702  
; Sequence 186702, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 186702  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_139604C.1.pep  
US-10-424-599-186702

Query Match 1.8%; Score 6; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 PIVEIL 294  
| | | | |  
Db 39 PIVEIL 44

RESULT 196  
US-10-424-599-188262  
; Sequence 188262, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188262  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141013C.1.pep  
US-10-424-599-188262

Query Match 1.8%; Score 6; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLAILE 29  
| | | | |  
Db 44 NLAILE 49

RESULT 197  
US-10-424-599-280571  
; Sequence 280571, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 280571  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

; LOCATION: (1)...(90)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_95379C.1.pep  
US-10-424-599-280571

Query Match 1.8%; Score 6; DB 12; Length 90;  
Best Local Similarity 100.0%; Pred. No. 8.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILXON 24  
| | | | |  
Db 61 KILXON 66

RESULT 198  
US-10-424-599-206794  
; Sequence 206794, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 206794  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_28762C.1.pep  
US-10-424-599-206794

Query Match 1.8%; Score 6; DB 12; Length 91;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 LGELIL 237  
| | | | |  
Db 43 LGELIL 48

RESULT 199  
US-09-864-408A-8758  
; Sequence 8758, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shimkets, Richard A.  
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enc  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8758  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-864-408A-8758

Query Match 1.8%; Score 6; DB 11; Length 92;  
Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 PAEIVK 19

```
Db      58 PABIVK 63
|||||
RESULT 200
US-10-424-599-157298
; Sequence 157298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 157298
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_113060C.1.pep
US-10-424-599-157298
Query Match      1.8%; Score 6; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      229 LKLLGE 234
|||||
Db      9 LKLLGE 14

RESULT 201
US-10-424-599-233814
; Sequence 233814, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 233814
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(92)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53159C.1.pep
US-10-424-599-233814
Query Match      1.8%; Score 6; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 IVKILK 22
|||||
Db      54 IVKILK 59

RESULT 202
US-10-424-599-161289
; Sequence 161289, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161289
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116662C.1.pep
US-10-424-599-161289
Query Match      1.8%; Score 6; DB 12; Length 93;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      293 ILLKNO 298
|||||
Db      41 ILLKNO 46

RESULT 203
US-10-424-599-256649
; Sequence 256649, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 256649
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73778C.1.pep
US-10-424-599-256649
Query Match      1.8%; Score 6; DB 12; Length 93;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      228 SLXLLG 233
|||||
Db      56 SLXLLG 61

RESULT 204
US-10-335-977-8937
; Sequence 8937, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
```







; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 194060  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_17262C.1.pap  
US-10-424-599-194060

Query Match 1.8%; Score 6; DB 12; Length 99;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPNLX 256  
Db 18 KPNLX 23

RESULT 213  
US-10-424-599-164813  
; Sequence 164813, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 164813  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_119844C.1.pap  
US-10-424-599-164813

Query Match 1.8%; Score 6; DB 12; Length 100;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 PTEAVA 67  
Db 78 PTEAVA 83

RESULT 214  
US-10-424-599-146233  
; Sequence 146233, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 146233  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Glycine max

; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_103066C.1.pap  
US-10-424-599-146233

Query Match 1.8%; Score 6; DB 12; Length 101;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295  
Db 16 IVEILL 21

RESULT 215  
US-10-094-749-1868  
; Sequence 1868, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOTIKO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1868  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1868

Query Match 1.8%; Score 6; DB 15; Length 101;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
Db 51 SLKLLG 56

RESULT 216  
US-09-925-302-852  
; Sequence 852, Application US/09925302  
; Patent No. US20020044941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925,302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 896  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 852  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (74)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (82)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (83)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (86)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (91)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (97)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-302-852

Query Match 1.8%; Score 6; DB 9; Length 102;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 198 VADFLE 203  
Db 8 VADFLE 13

RESULT 217  
US-09-864-761-36927  
; Sequence 36927, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 36927  
; LENGTH: 102  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL121778.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 20  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 75  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 36  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.9  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 43  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 18  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.9  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 34  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 23  
; OTHER INFORMATION: SWISSPROT HIT: P25444, EVALUE 1.00e-53  
; OTHER INFORMATION: EST\_HUMAN HIT: BF309937.1, EVALUE 1.00e-52  
US-09-864-761-36927

Query Match 1.8%; Score 6; DB 9; Length 102;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 277 FKVFVA 282  
Db 17 FKVFVA 22

RESULT 218  
US-10-424-599-249236  
; Sequence 249236, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; PRIOR FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 249236  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) ..(103)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_67090C.1.pap  
US-10-424-599-249236

Query Match 1.8%; Score 6; DB 12; Length 103;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TLIA DL 86  
|||||  
Db 73 TLIA DL 78

## RESULT 219

US-09-867-550-76  
; Sequence 76, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 76  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: VARIANT  
; LOCATION: (1)  
; OTHER INFORMATION: wherein Xaa may be any one of Leu or Ser or Trp or Pro or Gln or  
; NAME/KEY: VARIANT  
; LOCATION: (104)  
; OTHER INFORMATION: wherein Xaa may be any one of Ala or Pro or Ser or Thr  
US-09-867-550-76

Query Match 1.8%; Score 6; DB 9; Length 104;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295  
|||||  
Db 60 IVEILL 65

## RESULT 220

US-09-731-872-478  
; Sequence 478, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US3.REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 478  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-09-731-872-478

Query Match 1.8%; Score 6; DB 9; Length 104;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLLTRH 194  
|||||  
Db 54 DLLTRH 59

## RESULT 221

US-09-876-997-478  
; Sequence 478, Application US/09876997  
; Publication No. US20030152921A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78.US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 478  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-876-997-478

Query Match 1.8%; Score 6; DB 10; Length 104;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLLTRH 194  
|||||  
Db 54 DLLTRH 59

## RESULT 222

US-09-771-161A-177  
; Sequence 177, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771,161A  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 177  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-177

Query Match 1.8%; Score 6; DB 9; Length 105;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 TRHKVL 197  
|||||  
Db 27 TRHKVL 32

## RESULT 223

US-10-424-599-216376  
; Sequence 216376, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216376  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(105)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37416C.1.pap  
US-10-424-599-216376

Query Match 1.8%; Score 6; DB 12; Length 105;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 HEPLAK 157  
|||||  
Db 8 HEPLAK 13

## RESULT 224

US-10-263-828-83  
; Sequence 83, Application US/10263828  
; Publication No. US20030138905A1  
; GENERAL INFORMATION:  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; TITLE OF INVENTION: Compositions isolated from bovine  
; TITLE OF INVENTION: mammary gland and methods for their use.  
; FILE REFERENCE: 11000.1044UIcon  
; CURRENT APPLICATION NUMBER: US/10/263,828  
; CURRENT FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 136  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 83  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-263-828-83

Query Match 1.8%; Score 6; DB 14; Length 106;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KKTDKA 38  
|||||  
Db 52 KKTDKA 57

## RESULT 225

US-10-424-599-200671  
; Sequence 200671, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 200671  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_23231C.1.pap  
US-10-424-599-200671

Query Match 1.8%; Score 6; DB 12; Length 107;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KULGEL 235  
|||||  
Db 15 KULGEL 20

## RESULT 226

US-10-424-599-267548  
; Sequence 267548, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 267548  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_83619C.1.pap  
US-10-424-599-267548

Query Match 1.8%; Score 6; DB 12; Length 107;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 HKTQPI 290  
|||||  
Db 8 HKTQPI 13

## RESULT 227

US-10-425-114-70895  
; Sequence 70895, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.

```
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70895
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3959-004-D11_FLI.pep
US-10-425-114-70895

Query Match          1.8%; Score 6; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLLGEL 235
Db 79 KLLGEL 84

RESULT 228
US-10-424-599-178787
; Sequence 178787, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 178787
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(108)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_132461C.1.pep
US-10-424-599-178787

Query Match          1.8%; Score 6; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KKTDKA 38
Db 19 KKTDKA 24

RESULT 229
US-10-424-599-191237
; Sequence 191237, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
```

```
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 191237
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14706C.1.pep
US-10-424-599-191237

Query Match          1.8%; Score 6; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 QSLKLL 232
Db 44 QSLKLL 49

RESULT 230
US-09-864-408A-6796
; Sequence 6796, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enco
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6796
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-6796

Query Match          1.8%; Score 6; DB 11; Length 109;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82
Db 12 GLLVTL 17

RESULT 231
US-10-282-122A-43428
; Sequence 43428, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43428  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-282-122A-43428

Query Match 1.8%; Score 6; DB 12; Length 109;  
Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EKLQOS 219  
|||||  
DB 26 EKLQOS 31

RESULT 232  
US-10-282-122A-71923  
; Sequence 71923, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITFA 034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 71923  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Staphylococcus haemolyticus  
US-10-282-122A-71923

Query Match 1.8%; Score 6; DB 12; Length 109;  
Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ADFLEQ 204  
|||||  
DB 57 ADFLEQ 62

RESULT 233  
US-10-424-599-258409  
; Sequence 258409, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 258409  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(109)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_75368C.1.pap  
US-10-424-599-258409

Query Match 1.8%; Score 6; DB 12; Length 109;  
Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ILKDNL 25  
|||||  
DB 70 ILKDNL 75

RESULT 234  
US-10-425-114-46701  
; Sequence 46701, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114



; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46701  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700570435\_Fli.pep  
US-10-425-114-46701

Query Match  
Best Local Similarity 1.8%; Score 6; DB 12; Length 110;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLLA 84  
| | | | |  
DB 71 LVTLLA 76  
| | | | |

RESULT 235  
US-10-424-599-266911  
; Sequence 266911, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 266911  
; LENGTH: 111  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(111)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_83040C.1.pep  
US-10-424-599-266911

Query Match  
Best Local Similarity 1.8%; Score 6; DB 12; Length 111;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
| | | | |  
DB 29 VKILKD 34  
| | | | |

RESULT 236  
US-10-424-599-273101  
; Sequence 273101, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 273101  
; LENGTH: 111  
; TYPE: PRT

; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_88632C.1.pep  
US-10-424-599-273101

Query Match  
Best Local Similarity 1.8%; Score 6; DB 12; Length 111;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 ILEKQD 32  
| | | | |  
DB 83 ILEKQD 88  
| | | | |

RESULT 237  
US-10-424-599-181030  
; Sequence 181030, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 181030  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_134484C.1.pep  
US-10-424-599-181030

Query Match  
Best Local Similarity 1.8%; Score 6; DB 12; Length 113;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 LIDFEG 93  
| | | | |  
DB 57 LIDFEG 62  
| | | | |

RESULT 238  
US-10-424-599-188662  
; Sequence 188662, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188662  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(113)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141376C.1.pep  
US-10-424-599-188662

Query Match 1.8%; Score 6; DB 12; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DNAIL 28  
Db 10 DNAIL 15

RESULT 239  
US-10-424-599-235299  
; Sequence 235299, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 235299  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_54502C.1.pap  
US-10-424-599-235299

Query Match 1.8%; Score 6; DB 12; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 YEXLLQ 218  
Db 100 YEXLLQ 105

RESULT 240  
US-09-864-761-37175  
; Sequence 37175, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37175  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC007248.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.1  
; OTHER INFORMATION: SWISSPROT HIT: P13504, EVALUATE 1.00e-11  
; OTHER INFORMATION: EST\_HUMAN HIT: AV696742.1, EVALUATE 2.00e-26  
US-09-864-761-37175

Query Match 1.8%; Score 6; DB 9; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 QSLKLL 232  
Db 63 QSLKLL 68

RESULT 241  
US-09-765-272-68  
; Sequence 68, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-765-272-68

Query Match      1.8%  Score 6; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 FEGKDD 96
Db 19 FEGKDD 24

RESULT 242
US-09-939-980-347
; Sequence 347, Application US/09939980
; Patent No. US20020082234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; Burnham, Martin
; Hodgson, John
; Knowles, David
; Lonetto, Michael
; Nicholas, Richard
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
; TITLE OF INVENTION: No. US20020082234A1e1 Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090

```

```

; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 347:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 347:
US-09-939-980-347

Query Match      1.8%  Score 6; DB 9; Length 117;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23
Db 1 VKILKD 6

RESULT 243
US-10-425-114-42612
; Sequence 42612, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42612
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700380835_FLI.pep
US-10-425-114-42612

Query Match      1.8%  Score 6; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KILKDN 24
Db 102 KILKDN 107

RESULT 244
US-10-278-087A-50
; Sequence 50, Application US/10278087A
; Publication No. US20030138817A1
; GENERAL INFORMATION:
; APPLICANT: Shuji Hinuma
; Yasuaki Ito
; Ryo Fujii
; TITLE OF INVENTION: G Protein Coupled Receptor Protein,
; Production, And Use Thereof
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwards & Angell, LLP
; STREET: 101 Federal Street
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02209
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/278,087A  
FILING DATE: 31-Jan-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/461,436  
FILING DATE: 14-DEC-1999  
APPLICATION NUMBER: 09/038,572  
FILING DATE: 11-MAR-1998  
APPLICATION NUMBER: 08/513,974  
FILING DATE: 14-SEP-1995  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
APPLICATION NUMBER: 7-093989  
FILING DATE: 19-APR-1995  
APPLICATION NUMBER: 7-057186  
FILING DATE: 16-MAR-1995  
APPLICATION NUMBER: 7-007177  
FILING DATE: 20-JAN-1995  
APPLICATION NUMBER: 6-326611  
FILING DATE: 28-DEC-1994  
APPLICATION NUMBER: 6-270017  
FILING DATE: 02-NOV-1994  
APPLICATION NUMBER: 6-236357  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-236356  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-189274  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189273  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CONLIN, DAVID G.  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 45753 DIV3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-439-4444  
TELEFAX: 617-439-4170  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-10-278-087A-50

Query Match 1.8%; Score 6; DB 14; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 KDVTQI 100  
|||  
Db 45 KDVTQI 50

## RESULT 245

US-10-424-599-240972  
Sequence 240972, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21 (53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 240972  
LENGTH: 121  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)...(121)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59624C.1.pep  
US-10-424-599-240972

Query Match 1.8%; Score 6; DB 12; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 RCGIML 146  
|||  
Db 32 RCGIML 37

## RESULT 246

US-10-424-599-272118  
Sequence 272118, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21 (53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 272118  
LENGTH: 122  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_87741C.1.pep  
US-10-424-599-272118

Query Match 1.8%; Score 6; DB 12; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 KPEMLK 256  
|||  
Db 103 KPEMLK 108

## RESULT 247

US-09-893-737-26  
Sequence 26, Application US/09893737  
Patent No. US20020110855A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Presnell, Scott R.  
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
FILE REFERENCE: 00-41  
CURRENT APPLICATION NUMBER: US/09/893,737  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/215,446  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 329  
SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-893-737-26

Query Match 1.8%; Score 6; DB 9; Length 123;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 71  
Db 101 VAQLAQ 106

## RESULT 248

US-10-424-599-175359  
; Sequence 175359, Application US/10424599  
; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 175359

; LENGTH: 125

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_129368C.1.pap  
US-10-424-599-175359

Query Match 1.8%; Score 6; DB 12; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AQLAQE 72  
Db 79 AQLAQE 84

## RESULT 249

US-10-424-599-186519  
; Sequence 186519, Application US/10424599  
; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 186519

; LENGTH: 125

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1) .. (125)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13943C.1.pap

US-10-424-599-186519

Query Match 1.8%; Score 6; DB 12; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 LAKIIL 160  
Db 89 LAKIIL 94

## RESULT 250

US-10-424-599-243645  
; Sequence 243645, Application US/10424599  
; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 243645

; LENGTH: 125

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_62036C.1.pap  
US-10-424-599-243645

Query Match 1.8%; Score 6; DB 12; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 QIRDLLK 333  
Db 42 QIRDLLK 47

Search completed: April 12, 2004, 10:42:31  
Job time : 49 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2004, 10:33:58 ; Search time 22 seconds  
(without alignments)  
790.816 Million cell updates/sec

Title: US-10-025-730-1  
Perfect score: 337  
Sequence: 1 MKQMPLFSKHKNPAPKIVK.....FADEKNVLIQIRDLKKTAP 337

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pdp.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pdp.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pdp.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pdp.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pdp.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	337	100.0	337	3	US-09-190-965-1
2	337	100.0	337	4	US-09-470-253-1
3	23	6.8	341	3	US-09-190-965-3
4	23	6.8	341	4	US-09-470-253-3
5	15	4.5	339	3	US-09-190-965-4
6	15	4.5	339	4	US-09-470-253-4
7	14	4.2	377	3	US-09-190-965-5
8	14	4.2	377	4	US-09-470-253-5
9	8	2.4	923	4	US-09-345-473E-23
10	7	2.1	97	6	5196333-9
11	7	2.1	99	4	US-09-621-976-5453
12	7	2.1	99	4	US-09-866-108A-15755
13	7	2.1	205	1	US-08-450-944-5
14	7	2.1	205	5	PCT-US96-07709-5
15	7	2.1	221	1	US-08-450-944-2
16	7	2.1	221	4	US-09-489-039A-9612
17	7	2.1	221	5	PCT-US96-07709-2
18	7	2.1	276	4	US-09-252-991A-30691
19	7	2.1	298	4	US-09-543-681A-4229
20	7	2.1	341	4	US-09-489-039A-9238
21	7	2.1	341	4	US-09-489-039A-10987
22	7	2.1	341	4	US-09-489-039A-11246
23	7	2.1	341	4	US-09-489-039A-11246
24	7	2.1	341	4	US-09-328-352-4130
25	7	2.1	343	4	US-09-489-039A-11316
26	7	2.1	343	4	US-09-134-000C-6452
27	7	2.1	350	4	US-09-252-991A-22992

28	7	2.1	384	4	US-09-741-243C-8	Sequence 8, Appli
29	7	2.1	449	4	US-09-489-039A-8778	Sequence 8778, Ap
30	7	2.1	449	4	US-09-489-039A-9088	Sequence 9088, Ap
31	7	2.1	490	4	US-09-056-285A-10	Sequence 10, Appl
32	7	2.1	492	3	US-08-724-466B-2	Sequence 2, Appli
33	7	2.1	492	3	US-08-882-164D-2	Sequence 2, Appli
34	7	2.1	492	4	US-08-336-235A-1	Sequence 1, Appli
35	7	2.1	497	1	US-08-649-432-1	Sequence 1, Appli
36	7	2.1	497	5	PCT-US95-14024-1	Sequence 1, Appli
37	7	2.1	504	2	US-08-645-900A-1	Sequence 1, Appli
38	7	2.1	504	2	US-08-882-238A-1	Sequence 1, Appli
39	7	2.1	504	2	US-08-667-790A-1	Sequence 1, Appli
40	7	2.1	504	3	US-09-220-459-1	Sequence 1, Appli
41	7	2.1	504	3	US-08-938-669A-32	Sequence 1, Appli
42	7	2.1	504	3	US-08-546-568B-1	Sequence 32, Appl
43	7	2.1	504	3	US-08-822-999-3	Sequence 1, Appli
44	7	2.1	504	4	US-09-056-285A-8	Sequence 3, Appli
45	7	2.1	504	4	US-09-306-828-32	Sequence 8, Appli
46	7	2.1	510	4	US-09-489-039A-12958	Sequence 32, Appl
47	7	2.1	514	4	US-09-800-729-124	Sequence 12958, A
48	7	2.1	551	4	US-09-741-243C-2	Sequence 124, App
49	7	2.1	620	4	US-09-489-039A-12036	Sequence 2, Appli
50	7	2.1	1239	1	US-08-026-138B-3	Sequence 12036, A
51	7	2.1	1745	4	US-09-800-729-89	Sequence 3, Appli
52	7	2.1	3168	4	US-09-489-039A-14067	Sequence 89, Appl
53	7	2.1	3433	4	US-09-091-501B-10	Sequence 14067, A
54	6	1.8	15	2	US-08-432-871C-75	Sequence 10, Appl
55	6	1.8	15	2	US-08-432-871C-82	Sequence 75, Appl
56	6	1.8	15	4	US-09-270-956-75	Sequence 82, Appl
57	6	1.8	15	4	US-09-270-956-82	Sequence 75, Appl
58	6	1.8	22	1	US-07-965-663A-4	Sequence 82, Appl
59	6	1.8	38	1	US-08-033-873-8	Sequence 4, Appli
60	6	1.8	38	2	US-08-356-832-8	Sequence 8, Appli
61	6	1.8	38	3	US-08-988-705-8	Sequence 8, Appli
62	6	1.8	40	1	US-08-033-873-7	Sequence 8, Appli
63	6	1.8	40	1	US-08-033-873-9	Sequence 7, Appli
64	6	1.8	40	2	US-08-356-832-7	Sequence 9, Appli
65	6	1.8	40	2	US-08-356-832-9	Sequence 9, Appli
66	6	1.8	40	3	US-08-988-705-7	Sequence 7, Appli
67	6	1.8	40	3	US-08-988-705-9	Sequence 9, Appli
68	6	1.8	46	2	US-09-057-762-21	Sequence 21, Appl
69	6	1.8	46	3	US-08-326-119A-21	Sequence 21, Appl
70	6	1.8	52	4	US-09-621-976-6437	Sequence 6437, Ap
71	6	1.8	52	4	US-09-621-976-6438	Sequence 6438, Ap
72	6	1.8	60	4	US-09-543-681A-6353	Sequence 6353, Ap
73	6	1.8	61	4	US-09-107-532A-5451	Sequence 5451, Ap
74	6	1.8	63	4	US-09-673-395A-510	Sequence 510, Appl
75	6	1.8	64	3	US-09-091-814-25	Sequence 25, Appl
76	6	1.8	66	4	US-09-134-001C-4534	Sequence 4534, Ap
77	6	1.8	73	4	US-09-621-976-5578	Sequence 5578, Ap
78	6	1.8	82	4	US-09-858-664A-32	Sequence 32, Appl
79	6	1.8	82	4	US-10-274-978-33	Sequence 33, Appl
80	6	1.8	85	4	US-09-134-000C-6127	Sequence 6127, Ap
81	6	1.8	87	4	US-09-732-210-888	Sequence 888, App
82	6	1.8	93	4	US-09-107-532A-4435	Sequence 4435, Ap
83	6	1.8	100	4	US-09-621-976-4269	Sequence 4269, Ap
84	6	1.8	104	4	US-09-621-976-5842	Sequence 5842, Ap
85	6	1.8	114	4	US-09-252-991A-18411	Sequence 18411, A
86	6	1.8	117	3	US-08-961-083-68	Sequence 68, Appl
87	6	1.8	117	4	US-08-936-165A-347	Sequence 347, App
88	6	1.8	117	4	US-09-536-784-68	Sequence 68, Appl
89	6	1.8	119	3	US-08-513-974B-50	Sequence 50, Appl
90	6	1.8	119	3	US-08-513-974B-354	Sequence 354, App
91	6	1.8	119	3	US-08-513-974B-355	Sequence 355, App
92	6	1.8	119	4	US-09-461-436B-50	Sequence 50, Appl
93	6	1.8	128	4	US-09-134-000C-5393	Sequence 5393, Ap
94	6	1.8	128	4	US-09-134-000C-5393	Sequence 4, Appli
95	6	1.8	129	4	US-09-481-161-4	Sequence 481, Ap
96	6	1.8	131	4	US-09-328-352-4824	Sequence 4824, Ap
97	6	1.8	133	4	US-09-634-238-372	Sequence 372, App
98	6	1.8	133	4	US-09-138-452A-673	Sequence 673, App
99	6	1.8	134	4	US-09-732-210-804	Sequence 804, App
100	6	1.8	134	4	US-09-732-210-1265	Sequence 1265, Ap

101	6	1.8	135	2	US-08-383-621-5	Sequence 5, Appli	174	6	1.8	295	3	US-09-374-493-13	Sequence 13, Appli
102	6	1.8	136	3	US-08-459-906-5	Sequence 5, Appli	175	6	1.8	295	3	US-09-374-824-13	Sequence 13, Appli
103	6	1.8	137	4	US-09-149-476-554	Sequence 554, App	176	6	1.8	295	3	US-09-374-492-13	Sequence 13, Appli
104	6	1.8	138	4	US-09-489-039A-11504	Sequence 11504, A	177	6	1.8	295	4	US-09-785-343-13	Sequence 13, Appli
105	6	1.8	139	4	US-09-489-039A-13120	Sequence 13120, A	178	6	1.8	295	4	US-09-134-000C-3737	Sequence 3737, Ap
106	6	1.8	140	4	US-09-370-838-73	Sequence 73, Appli	179	6	1.8	299	2	US-08-923-856-1	Sequence 1, Appli
107	6	1.8	151	2	US-08-874-783-1	Sequence 1, Appli	180	6	1.8	299	3	US-09-216-294-1	Sequence 1, Appli
108	6	1.8	151	4	US-08-858-207A-423	Sequence 423, App	181	6	1.8	299	3	US-09-107-532A-4050	Sequence 4050, Ap
109	6	1.8	155	4	US-09-252-991A-30905	Sequence 30905, A	182	6	1.8	300	3	US-09-092-913-5	Sequence 5, Appli
110	6	1.8	160	4	US-09-134-001C-4854	Sequence 4854, Ap	183	6	1.8	300	4	US-09-976-594-7	Sequence 7, Appli
111	6	1.8	161	4	US-09-732-210-532	Sequence 532, App	184	6	1.8	303	4	US-09-252-991A-27266	Sequence 27266, A
112	6	1.8	162	4	US-09-732-210-531	Sequence 531, App	185	6	1.8	307	4	US-09-328-352-4994	Sequence 4994, Ap
113	6	1.8	163	4	US-09-328-352-5779	Sequence 5779, Ap	186	6	1.8	308	4	US-09-461-325-437	Sequence 437, App
114	6	1.8	170	4	US-09-252-991A-19902	Sequence 19902, A	187	6	1.8	308	4	US-09-252-991A-27921	Sequence 27921, A
115	6	1.8	174	4	US-09-149-476-526	Sequence 526, App	188	6	1.8	308	4	US-09-252-991A-29243	Sequence 29243, A
116	6	1.8	178	4	US-09-328-352-4965	Sequence 4965, Ap	189	6	1.8	308	4	US-09-328-352-7093	Sequence 7093, Ap
117	6	1.8	181	4	US-09-107-532A-5800	Sequence 5800, App	190	6	1.8	308	4	US-09-328-352-7158	Sequence 7158, Ap
118	6	1.8	181	4	US-09-461-325-196	Sequence 196, App	191	6	1.8	308	4	US-10-012-542-437	Sequence 437, App
119	6	1.8	183	4	US-09-198-452A-612	Sequence 612, App	192	6	1.8	313	4	US-09-655-908-12	Sequence 12, Appli
120	6	1.8	184	4	US-09-489-039A-7656	Sequence 7656, Ap	193	6	1.8	314	4	US-09-252-991A-22411	Sequence 22411, A
121	6	1.8	184	4	US-09-489-039A-12490	Sequence 12490, A	194	6	1.8	314	4	US-09-489-039A-8342	Sequence 8342, Ap
122	6	1.8	188	4	US-09-489-039A-8986	Sequence 8986, Ap	195	6	1.8	314	4	US-09-634-137-30	Sequence 30, Appli
123	6	1.8	190	4	US-09-107-532A-4452	Sequence 4452, Ap	196	6	1.8	315	1	US-08-571-758-12	Sequence 12, Appli
124	6	1.8	207	4	US-09-252-991A-22287	Sequence 22287, A	197	6	1.8	315	1	US-08-909-984A-12	Sequence 12, Appli
125	6	1.8	207	4	US-08-811-519-24	Sequence 24, Appli	198	6	1.8	315	1	US-09-107-532A-5596	Sequence 5596, Ap
126	6	1.8	209	4	US-09-134-001C-3080	Sequence 3080, Ap	199	6	1.8	317	4	US-09-134-001C-4204	Sequence 4204, Ap
127	6	1.8	212	4	US-09-391-741A-10	Sequence 10, Appli	200	6	1.8	320	4	US-09-848-294-4	Sequence 4, Appli
128	6	1.8	212	4	US-09-391-741A-24	Sequence 24, Appli	201	6	1.8	322	4	US-09-252-991A-18724	Sequence 18724, A
129	6	1.8	212	4	US-09-391-741A-34	Sequence 34, Appli	202	6	1.8	322	4	US-08-475-612A-4	Sequence 4, Appli
130	6	1.8	214	4	US-09-391-741A-16	Sequence 16, Appli	203	6	1.8	323	2	US-08-475-612A-4	Sequence 4, Appli
131	6	1.8	214	4	US-09-391-741A-26	Sequence 26, Appli	204	6	1.8	323	3	US-09-191-359-4	Sequence 12, Appli
132	6	1.8	214	4	US-09-328-352-6571	Sequence 6571, Ap	205	6	1.8	323	3	US-09-097-231-12	Sequence 12, Appli
133	6	1.8	218	4	US-09-634-238-341	Sequence 341, App	206	6	1.8	323	3	US-09-353-099-12	Sequence 12, Appli
134	6	1.8	218	4	US-09-328-352-6222	Sequence 6222, Ap	207	6	1.8	323	4	US-09-252-991A-26419	Sequence 26419, A
135	6	1.8	220	3	US-09-247-373B-50	Sequence 50, Appli	208	6	1.8	323	4	US-09-134-000C-5535	Sequence 5535, Ap
136	6	1.8	222	4	US-09-134-001C-3053	Sequence 3053, Ap	209	6	1.8	323	4	US-09-709-066-2	Sequence 2, Appli
137	6	1.8	227	4	US-09-489-039A-9910	Sequence 9910, Ap	210	6	1.8	324	4	US-09-370-838-45	Sequence 45, Appli
138	6	1.8	230	4	US-09-107-532A-5328	Sequence 5328, Ap	211	6	1.8	331	4	US-09-489-039A-10843	Sequence 23, Appli
139	6	1.8	231	4	US-09-543-681A-6078	Sequence 6078, Ap	212	6	1.8	332	4	US-09-489-039A-9330	Sequence 9330, Ap
140	6	1.8	231	4	US-09-489-039A-10973	Sequence 10973, A	213	6	1.8	332	4	US-09-134-000C-5207	Sequence 5207, Ap
141	6	1.8	232	1	US-08-468-036-7	Sequence 7, Appli	214	6	1.8	334	4	US-09-674-681A-5285	Sequence 5285, Ap
142	6	1.8	232	2	US-08-376-843-7	Sequence 4013, Ap	215	6	1.8	336	4	US-09-107-532A-6417	Sequence 6417, Ap
143	6	1.8	232	4	US-09-134-001C-4013	Sequence 4013, Ap	216	6	1.8	337	2	US-09-057-762-2	Sequence 2, Appli
144	6	1.8	242	4	US-09-198-452A-182	Sequence 182, App	217	6	1.8	337	3	US-08-328-119A-2	Sequence 2, Appli
145	6	1.8	243	3	US-09-252-991A-29870	Sequence 29870, A	218	6	1.8	338	4	US-09-198-452A-642	Sequence 642, App
146	6	1.8	247	4	US-09-100-804-13	Sequence 13, Appli	219	6	1.8	338	4	US-09-614-912-150	Sequence 150, App
147	6	1.8	247	4	US-09-489-039A-7921	Sequence 7921, Ap	220	6	1.8	346	2	US-08-276-151-5	Sequence 5, Appli
148	6	1.8	249	4	US-09-489-039A-8345	Sequence 8345, Ap	221	6	1.8	346	2	US-08-476-254-2	Sequence 2, Appli
149	6	1.8	254	4	US-09-252-991A-25583	Sequence 25583, A	222	6	1.8	346	2	US-08-476-254-10	Sequence 10, Appli
150	6	1.8	255	4	US-09-252-991A-19444	Sequence 19444, A	223	6	1.8	346	6	US-09-153-599A-10	Sequence 10, Appli
151	6	1.8	257	4	US-08-311-731A-141	Sequence 141, App	224	6	1.8	346	6	5474933-2	Patent No. 5474933
152	6	1.8	257	4	US-09-634-955B-23	Sequence 23, Appli	225	6	1.8	346	6	US-09-252-991A-31854	Sequence 31854, A
153	6	1.8	263	4	US-09-252-991A-25534	Sequence 25534, A	226	6	1.8	347	6	US-09-489-039A-7292	Sequence 7292, Ap
154	6	1.8	263	4	US-09-328-352-4853	Sequence 4853, Ap	227	6	1.8	347	6	US-09-198-452A-702	Sequence 702, App
155	6	1.8	265	4	US-09-198-452A-23	Sequence 23, Appli	228	6	1.8	351	4	US-09-134-000C-9887	Sequence 9887, Ap
156	6	1.8	267	2	US-07-857-224B-42	Sequence 42, Appli	229	6	1.8	352	4	US-08-978-182-4	Sequence 4914, Ap
157	6	1.8	269	4	US-09-252-991A-31810	Sequence 31810, A	230	6	1.8	352	4	US-09-205-681-4	Sequence 4, Appli
158	6	1.8	270	3	US-09-101-146-62	Sequence 62, Appli	231	6	1.8	359	4	US-09-634-955B-21	Sequence 21, Appli
159	6	1.8	272	4	US-09-134-001C-2864	Sequence 2864, Ap	232	6	1.8	362	2	US-08-252-991A-21741	Sequence 21741, A
160	6	1.8	274	4	US-09-489-039A-13748	Sequence 13748, A	233	6	1.8	362	3	US-09-080-897-6	Sequence 6, Appli
161	6	1.8	274	4	US-09-134-001C-3921	Sequence 3921, Ap	234	6	1.8	362	3	US-09-323-735-6	Sequence 6, Appli
162	6	1.8	279	4	US-09-328-352-4299	Sequence 4299, Ap	235	6	1.8	362	4	US-09-594-669-6	Sequence 6, Appli
163	6	1.8	281	4	US-09-328-352-7494	Sequence 7494, Ap	236	6	1.8	362	4	US-10-112-432-6	Sequence 6, Appli
164	6	1.8	282	4	US-09-540-236-3538	Sequence 4229, Ap	237	6	1.8	363	4	US-09-252-991A-27806	Sequence 27806, A
165	6	1.8	283	3	US-09-009-913-339	Sequence 339, App	238	6	1.8	363	4	US-09-489-039A-13446	Sequence 13446, A
166	6	1.8	284	4	US-09-071-035-426	Sequence 426, App	239	6	1.8	370	4	US-09-107-532A-5920	Sequence 5920, Ap
167	6	1.8	287	4	US-09-292-858B-25	Sequence 25, Appli	240	6	1.8				
168	6	1.8	287	4	US-09-543-681A-6438	Sequence 6438, Ap	241	6	1.8				
169	6	1.8	294	4	US-09-134-000C-4481	Sequence 4481, Ap	242	6	1.8				
170	6	1.8	295	3	US-09-150-141-13	Sequence 13, Appli	243	6	1.8				
171	6	1.8	295	3			244	6	1.8				
172	6	1.8					245	6	1.8				
173	6	1.8					246	6	1.8				

247 1.8 371 1 US-08-746-789A-2  
248 1.8 371 4 US-09-570-593-5  
249 1.8 371 4 US-09-543-681A-6429  
250 1.8 373 4 US-09-489-039A-8389  
251 1.8 382 2 US-08-477-451-28  
252 1.8 382 4 US-09-328-352-7860  
253 1.8 383 1 US-08-486-037B-2  
254 1.8 385 4 US-09-489-039A-8660  
255 1.8 391 4 US-09-594-669-4  
256 1.8 391 4 US-10-112-432-4  
257 1.8 396 1 US-08-769-309A-15  
258 1.8 396 3 US-08-994-570-15  
259 1.8 397 4 US-09-328-352-7457  
260 1.8 398 4 US-09-328-352-5257  
261 1.8 402 4 US-09-252-991A-17911  
262 1.8 402 4 US-09-107-532A-5360  
263 1.8 404 1 US-08-696-139-2  
264 1.8 405 4 US-09-594-669-12  
265 1.8 405 4 US-10-112-432-12  
266 1.8 409 4 US-09-564-559B-7  
267 1.8 409 4 US-09-564-559B-8  
268 1.8 409 4 US-09-564-559B-9  
269 1.8 409 4 US-09-564-559B-10  
270 1.8 410 3 US-09-023-809B-2  
271 1.8 410 4 US-09-328-352-5085  
272 1.8 414 4 US-09-252-991A-21182  
273 1.8 414 4 US-09-489-039A-11717  
274 1.8 417 4 US-09-134-001C-5006  
275 1.8 421 4 US-09-585-173B-52  
276 1.8 425 2 US-08-986-963-2  
277 1.8 427 4 US-09-107-532A-6633  
278 1.8 427 4 US-09-489-039A-11613  
279 1.8 429 4 US-09-328-352-4643  
280 1.8 430 3 US-09-007-484-2  
281 1.8 430 3 US-09-309-682-2  
282 1.8 430 4 US-09-594-669-2  
283 1.8 430 4 US-09-134-000C-4980  
284 1.8 430 4 US-10-112-432-2  
285 1.8 431 4 US-09-328-352-4593  
286 1.8 432 4 US-09-489-039A-9016  
287 1.8 434 4 US-09-594-669-10  
288 1.8 434 4 US-10-112-432-10  
289 1.8 436 1 US-08-080-255-8  
290 1.8 436 3 US-08-465-713-8  
291 1.8 436 5 PCT-US93-05857-8  
292 1.8 437 4 US-09-328-352-6841  
293 1.8 441 4 US-09-107-532A-4173  
294 1.8 445 4 US-09-414-189-1  
295 1.8 450 3 US-09-191-879-2  
296 1.8 451 4 US-09-489-039A-11654  
297 1.8 453 1 US-08-769-309A-14  
298 1.8 453 3 US-08-994-570-14  
299 1.8 454 4 US-09-252-991A-30013  
300 1.8 457 3 US-09-416-213-2

ALIGNMENTS

RESULT 1  
US-09-190-965-1  
; Sequence 1, Application US/09190965  
; Patent No. 6071721  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/190,965  
; CURRENT FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 3734805  
US-09-190-965-1

Query Match 100.0%; Score 337; DB 3; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKTKASEEVSKSLOAMKEILCGTNEKE 60  
DB 1 MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKTKASEEVSKSLOAMKEILCGTNEKE 60  
QY 61 PPTAVAGLAQELYSGLLVTLIADQLIDFEGKDVQIIFNNILRRQIGTRSPVVEIS 120  
DB 61 PPTAVAGLAQELYSGLLVTLIADQLIDFEGKDVQIIFNNILRRQIGTRSPVVEIS 120  
QY 121 APHILFLLKGYAPQIALRCGIMLRCEIRHEPLAKILFNSNFRDPFKYVELSTFDIA 180  
DB 121 APHILFLLKGYAPQIALRCGIMLRCEIRHEPLAKILFNSNFRDPFKYVELSTFDIA 180  
QY 181 SDAFATFKDLTRHKVLVADFLQNYDTIFEDYEKLGQSENVYTKRQSLKLGELIDRH 240  
DB 181 SDAFATFKDLTRHKVLVADFLQNYDTIFEDYEKLGQSENVYTKRQSLKLGELIDRH 240  
QY 241 NFATMTKYSKPNKLMNNILRRDKSPNIOFEAFHVKFVSPHKTQPIVEILLKNQPK 300  
DB 241 NFATMTKYSKPNKLMNNILRRDKSPNIOFEAFHVKFVSPHKTQPIVEILLKNQPK 300  
QY 301 LIEFLSFQKERTDDEQFADEKNVLIKOIRDLKKTAP 337  
DB 301 LIEFLSFQKERTDDEQFADEKNVLIKOIRDLKKTAP 337

RESULT 2  
US-09-470-253-1  
; Sequence 1, Application US/09470253  
; Patent No. 6365371  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470,253  
; CURRENT FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 3734805  
US-09-470-253-1

Query Match 100.0%; Score 337; DB 4; Length 337;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKTKASEEVSKSLOAMKEILCGTNEKE 60  
DB 1 MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKTKASEEVSKSLOAMKEILCGTNEKE 60  
QY 61 PPTAVAGLAQELYSGLLVTLIADQLIDFEGKDVQIIFNNILRRQIGTRSPVVEIS 120



Db 61 PPTAAVAQAQELSSGLLVTLADIQLDFEKKDVTQIFNNILRRQIGTRSPVVEYIS 120  
QY 121 ANPHILFMLLKGYEAPQIALRCGIMLRGICIRHEPILAKIILFSNQPRDFPKYVELSTFDIA 180  
Db 121 ANPHILFMLLKGYEAPQIALRCGIMLRGICIRHEPILAKIILFSNQPRDFPKYVELSTFDIA 180  
QY 181 SDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLQSENVVTKRQSLKLGELILDRH 240  
Db 181 SDAFATFKDLLTRHKVLVADFLQNYDTTFEDYEKLQSENVVTKRQSLKLGELILDRH 240  
QY 241 NFAIMTKYISKPENLKLMMNLLRDKSPNIQPFAPFVKFVSPHKTQPIVEILLKKNQPK 300  
Db 241 NFAIMTKYISKPENLKLMMNLLRDKSPNIQPFAPFVKFVSPHKTQPIVEILLKKNQPK 300  
QY 301 LIEFLSSFOKERTDDEQFADEKNYLKQIRDLKKTAP 337  
Db 301 LIEFLSSFOKERTDDEQFADEKNYLKQIRDLKKTAP 337

RESULT 3  
US-09-190-965-3  
; Sequence 3, Application US/09190965  
; Patent No. 6071721  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/190,965  
; CURRENT FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; FEATURE: -  
; OTHER INFORMATION: g262934  
US-09-190-965-3

Query Match 6.8%; Score 23; DB 3; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1.3e-14;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKLMMNLLRDKS 266  
Db 245 IMTKYISKPENLKLMMNLLRDKS 267

RESULT 4  
US-09-470-253-3  
; Sequence 3, Application US/09470253  
; Patent No. 6365371  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470,253  
; CURRENT FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Mus sp.  
; FEATURE: -  
; OTHER INFORMATION: g262934

; OTHER INFORMATION: g262934  
US-09-470-253-3

Query Match 6.8%; Score 23; DB 4; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1.3e-14;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYISKPENLKLMMNLLRDKS 266  
Db 245 IMTKYISKPENLKLMMNLLRDKS 267

RESULT 5  
US-09-190-965-4  
; Sequence 4, Application US/09190965  
; Patent No. 6071721  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/190,965  
; CURRENT FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 339  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE: -  
; OTHER INFORMATION: g1794137  
US-09-190-965-4

Query Match 4.5%; Score 15; DB 3; Length 339;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPVVEYI 119  
Db 103 LRRQIGTRSPVVEYI 117

RESULT 6  
US-09-470-253-4  
; Sequence 4, Application US/09470253  
; Patent No. 6365371  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470,253  
; CURRENT FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/190,965  
; PRIOR FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 339  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE: -  
; OTHER INFORMATION: g1794137  
US-09-470-253-4

Query Match 4.5%; Score 15; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPVEYI 119  
Db 103 LRRQIGTRSPVEYI 117

## RESULT 7

US-09-190-965-5  
; Sequence 5, Application US/09190965  
; Patent No. 6071721  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/190,965  
; CURRENT FILING DATE: 1998-11-13  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
; FEATURE: -  
; OTHER INFORMATION: g1255838  
US-09-190-965-5

Query Match 4.2%; Score 14; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPVEYI 118  
Db 119 LRRQIGTRSPVEYI 132

## RESULT 8

US-09-470-253-5  
; Sequence 5, Application US/09470253  
; Patent No. 6365371  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina A.  
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN  
; FILE REFERENCE: PF-0635 US  
; CURRENT APPLICATION NUMBER: US/09/470,253  
; CURRENT FILING DATE: 1999-12-22  
; PRIOR APPLICATION NUMBER: 09/190,965  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
; FEATURE: -  
; OTHER INFORMATION: g1255838  
US-09-470-253-5

Query Match 4.2%; Score 14; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIGTRSPVEYI 118  
Db 119 LRRQIGTRSPVEYI 132

## RESULT 9

US-09-345-473E-23

; Sequence 23, Application US/09345473E  
; Patent No. 6558903  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin  
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof  
; FILE REFERENCE: 35800/183781  
; CURRENT APPLICATION NUMBER: US/09/345,473E  
; CURRENT FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 923  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-345-473E-23

Query Match 2.4%; Score 8; DB 4; Length 923;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLVTL 82  
Db 449 SSGLLVTL 456

## RESULT 10

5196333-9  
; Patent No. 5196333  
; APPLICANT: CHALFIE, MARIN; WOLINSKY, EVE; DRISCOLL, MONICA  
; TITLE OF INVENTION: DNA SEQUENCES INVOLVED IN NEURONAL  
; DEGENERATION, MULTICELLULAR ORGANISMS CONTAINING SAME AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 11  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/530,968  
; FILING DATE: 30-MAY-1990  
; SEQ ID NO: 9;  
; LENGTH: 97  
5196333-9

Query Match 2.1%; Score 7; DB 6; Length 97;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLIAD 85  
Db 45 LVTLIAD 51

## RESULT 11

US-09-621-976-5453  
; Sequence 5453, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5453  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -73...-1  
US-09-621-976-5453

Query Match 2.1%; Score 7; DB 4; Length 99;

Mon Apr 12 10:49:27 2004

us-10-025-730-1.oligo.ra1

Best Local Similarity 100.0%; Pred. No. 31; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 59 KEPTTEA 65  
DB 87 KEPTTEA 93

RESULT 12  
US-08-866-108A-15755  
; Sequence 15755, Application US/09866108A  
; Patent No. 6686188  
; GENERAL INFORMATION:  
; APPLICANT: GU, Yizhong  
; APPLICANT: JI, Yonggang  
; APPLICANT: PENN, Sharon G.  
; APPLICANT: HANZEL, David K.  
; APPLICANT: RANK, David R.  
; APPLICANT: CHEN, Wensheng  
; APPLICANT: SHANNON, Mark  
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE  
; FILE REFERENCE: A6MICA-7  
; CURRENT APPLICATION NUMBER: US/09/866,108A  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 15755  
; SOFTWARE: A6MICA Sequence Listing Engine  
; Patent No. 6686188  
; SEQ ID NO 15755  
; LENGTH: 99  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-866-108A-15755

Query Match 2.1%; Score 7; DB 4; Length 99;  
Best Local Similarity 100.0%; Pred. No. 31; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 260 NLRDKS 266  
DB 63 NLRDKS 69

RESULT 13  
US-08-450-944-5  
; Sequence 5, Application US/08450944  
; Patent No. 5789194  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN  
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS  
; NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,944  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: 2618-30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 205 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-450-944-5

Query Match 2.1%; Score 7; DB 1; Length 205;  
Best Local Similarity 100.0%; Pred. No. 61; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 330 RDLKTA 336  
DB 93 RDLKTA 99

RESULT 14  
PCT-US96-07709-5  
; Sequence 5, Application PC/TUS9607709  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN  
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS  
; NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/07709  
FILING DATE: 23-MAY-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-30-PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 5:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-07709-5

Query Match          2.1%; Score 7; DB 5; Length 205;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 RDLKTKTA 336
DB 93 RDLKTKTA 99

RESULT 15
US-08-450-944-2
; Sequence 2, Application US/08450944
; Patent No. 5789194
; GENERAL INFORMATION:
; APPLICANT: Tripp, Cynthia A.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,944
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: 2618-30
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-944-2

Query Match          2.1%; Score 7; DB 1; Length 221;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 RDLKTKTA 336
DB 109 RDLKTKTA 115

RESULT 16
US-09-489-039A-9612
; Sequence 9612, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Berton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
```

```
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9612
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9612

Query Match          2.1%; Score 7; DB 4; Length 221;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AEIVKIL 21
DB 117 AEIVKIL 123

RESULT 17
PCT-US96-07709-2
; Sequence 2, Application PC/TUS9607709
; GENERAL INFORMATION:
; APPLICANT: Tripp, Cynthia A.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07709
; FILING DATE: 23-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-30-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-07709-2

Query Match          2.1%; Score 7; DB 5; Length 221;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 RDLKTKTA 336
DB 109 RDLKTKTA 115

RESULT 18
US-09-252-991A-30691
```

```
; Sequence 30691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30691
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30691

Query Match      2.1%; Score 7; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      65 AVAQLAQ 71
Db      65 AVAQLAQ 71

RESULT 19
US-09-543-681A-4229
; Sequence 4229, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4229
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4229

Query Match      2.1%; Score 7; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      87 QLIDFEG 93
Db      55 QLIDFEG 61

RESULT 20
US-09-489-039A-9238
; Sequence 9238, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9238
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9238

Query Match      2.1%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      299 PKLIEFL 305
Db      43 PKLIEFL 49

RESULT 21
US-09-489-039A-10987
; Sequence 10987, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10987
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10987

Query Match      2.1%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      299 PKLIEFL 305
Db      43 PKLIEFL 49

RESULT 22
US-09-489-039A-11246
; Sequence 11246, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11246
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11246

Query Match      2.1%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      299 PKLIEFL 305
Db      43 PKLIEFL 49

RESULT 23
US-09-489-039A-11246
; Sequence 11246, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9238
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11246

Query Match      2.1%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      299 PKLIEFL 305
Db      43 PKLIEFL 49

RESULT 23
```

US-09-489-039A-11507  
; Sequence 11507, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11507  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11507

Query Match 2.1%; Score 7; DB 4; Length 341;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
Db 43 PKLIEFL 49

RESULT 24  
US-09-328-352-4130  
; Sequence 4130, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4130  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4130

Query Match 2.1%; Score 7; DB 4; Length 343;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
Db 45 PKLIEFL 51

RESULT 25  
US-09-489-039A-11316  
; Sequence 11316, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11316  
; LENGTH: 343  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-11316

Query Match 2.1%; Score 7; DB 4; Length 343;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 PKLIEFL 305  
Db 45 PKLIEFL 51

RESULT 26  
US-09-134-000C-6452  
; Sequence 6452, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lythn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6452  
; LENGTH: 345  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6452

Query Match 2.1%; Score 7; DB 4; Length 345;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 ALLEKQD 32  
Db 21 ALLEKQD 27

RESULT 27  
US-09-252-991A-22992  
; Sequence 22992, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22992  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22992

Query Match 2.1%; Score 7; DB 4; Length 350;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 VSKSLQA 48  
Db 4 VSKSLQA 10

RESULT 28

US-09-741-243C-8  
; Sequence 8, Application US/09741243C  
; Patent No. 6399352  
; GENERAL INFORMATION:  
; APPLICANT: Crawford Jr., John Milton  
; APPLICANT: Rice, John  
; APPLICANT: Sevala, Veeresh  
; APPLICANT: Stewart, Sandy  
; TITLE OF INVENTION: NOVEL PLANT PORPHOBILINOGEN AND FUSION  
; TITLE OF INVENTION: PROTEIN THEREOF  
; FILE REFERENCE: 2022US  
; CURRENT APPLICATION NUMBER: US/09/741,243C  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/171,785  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
US-09-741-243C-8  
Query Match 2.1%; Score 7; DB 4; Length 384;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 261 LLRDKSP 267  
DB 316 LLRDKSP 322  
RESULT 29  
US-09-489-039A-8778  
; Sequence 8778, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709-2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8778  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8778  
Query Match 2.1%; Score 7; DB 4; Length 449;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 240 HNFAMT 246  
DB 127 HNFAMT 133  
RESULT 30  
US-09-489-039A-9088  
; Sequence 9088, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709-2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747

US-09-489-039A-9088  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9088  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9088  
Query Match 2.1%; Score 7; DB 4; Length 449;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 174 LSTFDIA 180  
DB 171 LSTFDIA 177  
RESULT 31  
US-09-056-285A-10  
; Sequence 10, Application US/09056285A  
; Patent No. 6403307  
; GENERAL INFORMATION:  
; APPLICANT: Stone, Edwin M.  
; APPLICANT: Sheffield, Val C.  
; APPLICANT: Alward, Wallace L.M.  
; APPLICANT: Fingert, John  
; TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,285A  
; FILING DATE: 07-Apr-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arnold, Beth E.  
; REGISTRATION NUMBER: 35,430  
; REFERENCE/DOCKET NUMBER: UTA-010.28  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 490 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-056-285A-10  
Query Match 2.1%; Score 7; DB 4; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 260 NLLRDKS 266  
DB 130 NLLRDKS 136  
RESULT 32  
US-08-724-466B-2  
; Sequence 2, Application US/08724466B  
; Patent No. 6063606  
; GENERAL INFORMATION:

APPLICANT: Petkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
ZIP: M5L 1A9  
COUNTRY: Canada  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,466B  
FILING DATE: October 1, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-466B-2

Query Match 2.1%; Score 7; DB 3; Length 492;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 SLOWAKE 51  
DB 282 SLOWAKE 288

RESULT 33  
US-08-882-164D-2  
Sequence 2, Application US/08882164D  
Patent No. 6306624  
GENERAL INFORMATION:  
APPLICANT: Petkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5L 1A9  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,164D  
FILING DATE: June 25, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
APPLICATION NUMBER: 08/724,466

FILING DATE: October 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-882-164D-2

Query Match 2.1%; Score 7; DB 4; Length 492;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 SLOWAKE 51  
DB 282 SLOWAKE 288

RESULT 34  
US-08-336-235A-1  
Sequence 1, Application US/08336235A  
Patent No. 5606043  
GENERAL INFORMATION:  
APPLICANT: NGUYEN, THAI D.  
APPLICANT: POLANSKY, JON R.  
APPLICANT: HUANG, WEIDONG  
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS OF GLAUCOMA  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWREY & SIMON  
STREET: 1299 PENNSYLVANIA AVE., N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: US  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,235A  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: AUERBACH, JEFFREY I  
REGISTRATION NUMBER: 32,680  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 383-7451  
TELEFAX: (202) 383-6610  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
IMMEDIATE SOURCE:  
CLONE: TIGR  
US-08-336-235A-1

Query Match 2.1%; Score 7; DB 1; Length 497;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266





```
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: No. 5849879e
US-08-645-900A-1

Query Match      2.1%; Score 7; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLLRDKS 266
DB      144 NLLRDKS 150

RESULT 38
US-08-882-238A-1
/ Sequence 1, Application US/08882238A
/ Patent No. 5854415
/ GENERAL INFORMATION:
/ APPLICANT: NGUYEN, THAI D.
/ APPLICANT: POLANSKY, JON R.
/ APPLICANT: HUANG, WEIDONG
/ TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS OF
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: HOWREY & SIMON
/ STREET: 1299 PENNSYLVANIA AVE., N.W.
/ CITY: WASHINGTON
/ STATE: D.C.
/ COUNTRY: US
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/882,238A
/ FILING DATE: 25 JUNE 1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MARSH, DAVID
/ REGISTRATION NUMBER: 41,408
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 383-7451
/ TELEFAX: (202) 383-6610
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 504 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: No. 5854415e
US-08-882-238A-1

Query Match      2.1%; Score 7; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLLRDKS 266
DB      144 NLLRDKS 150

RESULT 39
US-08-667-790A-1
/ Sequence 1, Application US/08667790A
/ Patent No. 5861497
/ GENERAL INFORMATION:
/ APPLICANT: NGUYEN, THAI D.
/ APPLICANT: POLANSKY, JON R.
```

```
/ APPLICANT: HUANG, WEIDONG
/ TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS OF
/ TITLE OF INVENTION: GLAUCOMA
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: HOWREY & SIMON
/ STREET: 1299 PENNSYLVANIA AVE., N.W.
/ CITY: WASHINGTON
/ STATE: D.C.
/ COUNTRY: US
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/667,790A
/ FILING DATE: 21 JUNE 1996
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MARSH, DAVID
/ REGISTRATION NUMBER: 41,408
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 383-7451
/ TELEFAX: (202) 383-6610
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 504 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: No. 5861497e
US-08-667-790A-1

Query Match      2.1%; Score 7; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      260 NLLRDKS 266
DB      144 NLLRDKS 150

RESULT 40
US-09-220-459-1
/ Sequence 1, Application US/09220459
/ Patent No. 6150161
/ GENERAL INFORMATION:
/ APPLICANT: Nguyen, Thai D
/ APPLICANT: Polansky, Jon R
/ APPLICANT: Huang, Weidong
/ TITLE OF INVENTION: Methods for the Diagnosis of Glaucoma
/ FILE REFERENCE: 07425.0056
/ CURRENT APPLICATION NUMBER: US/09/220,459
/ CURRENT FILING DATE: 1998-12-24
/ EARLIER APPLICATION NUMBER: 08/882,238
/ EARLIER FILING DATE: 1997-06-25
/ EARLIER APPLICATION NUMBER: 08/649,432
/ EARLIER FILING DATE: 1996-05-17
/ EARLIER APPLICATION NUMBER: 08/546,568
/ EARLIER FILING DATE: 1995-10-20
/ EARLIER APPLICATION NUMBER: 08/336,235
/ EARLIER FILING DATE: 1994-11-03
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 504
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-220-459-1

Query Match      2.1%; Score 7; DB 3; Length 504;
```

```
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266
Db 144 NLLRDKS 150

RESULT 41
US-08-938-669A-32
; Sequence 32, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; US-08-938-669A-32

Query Match 2.1%; Score 7; DB 3; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266
Db 144 NLLRDKS 150

RESULT 42
US-08-546-568B-1
; Sequence 1, Application US/08546568B
; Patent No. 6248867
; GENERAL INFORMATION:
; APPLICANT: NGUYEN, THAI D.
; APPLICANT: POLANSKY, JON R.
; APPLICANT: HUANG, WEIDONG
```

```
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS OF
; TITLE OF INVENTION: GLAUCOMA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWREY & SIMON
; STREET: 1299 PENNSYLVANIA AVE., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/546,568B
; APPLICATION NUMBER: US/08/546,568B
; FILING DATE: 20 October 1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MARSH, DAVID
; REGISTRATION NUMBER: 41,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 383-7451
; TELEFAX: (202) 383-6610
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6248867e
; US-08-546-568B-1

Query Match 2.1%; Score 7; DB 3; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLLRDKS 266
Db 144 NLLRDKS 150

RESULT 43
US-08-822-999-3
; Sequence 3, Application US/08822999
; Patent No. 6271026
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; TITLE OF INVENTION: GLAUCOMA COMPOSITIONS AND THERAPEUTIC
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,999
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/791,347
; FILING DATE: 30-JAN-1997
```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/748,479  
;; FILING DATE: 08-NOV-1996  
;; PRIOR APPLICATION DATA: US 08/234,218  
;; APPLICATION NUMBER: US 08/234,218  
;; FILING DATE: 28-APR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Arnold, Beth E.  
;; REGISTRATION NUMBER: 35,430  
;; REFERENCE/DOCKET NUMBER: UIA-010.27  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-832-1000  
;; TELEFAX: 617-832-7000  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 504 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-822-999-3

Query Match 2.1%; Score 7; DB 3; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLRDKS 266  
|||  
Db 144 NLRDKS 150

## RESULT 44

US-09-056-285A-8  
; Sequence 8, Application US/09056285A  
; Patent No. 6403307

## GENERAL INFORMATION:

;; APPLICANT: Stone, Edwin M.  
;; Sheffield, Val C.  
;; Alward, Wallace L.M.  
;; Fingert, John

## TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS

NUMBER OF SEQUENCES: 43

## CORRESPONDENCE ADDRESSES:

;; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
;; STREET: One Post Office Square  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109-2170

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/056,285A  
;; FILING DATE: 07-Apr-1998

## ATTORNEY/AGENT INFORMATION:

;; NAME: Arnold, Beth E.  
;; REGISTRATION NUMBER: 35,430  
;; REFERENCE/DOCKET NUMBER: UIA-010.28  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-832-1000  
;; TELEFAX: 617-832-7000

## INFORMATION FOR SEQ ID NO: 8:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 504 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein

;; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-056-285A-8

## Query Match

2.1%; Score 7; DB 4; Length 504;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLRDKS 266  
|||  
Db 144 NLRDKS 150

## RESULT 45

US-09-306-828-32  
; Sequence 32, Application US/09306828  
; Patent No. 6475724

## GENERAL INFORMATION:

;; APPLICANT: Nguyen, Thai D.  
;; APPLICANT: Polansky, Jon R.  
;; APPLICANT: Chen, Pu  
;; APPLICANT: Chen, Hua

;; TITLE OF INVENTION: Nucleic Acids, Kits, And Methods For The Diagnosis, Prognosis And  
;; CURRENT APPLICATION NUMBER: US/09/306,828

;; CURRENT FILING DATE: 1999-05-07

;; EARLIER APPLICATION NUMBER: US 09/227,881

;; EARLIER FILING DATE: 1999-01-11

;; NUMBER OF SEQ ID NOS: 38

;; SOFTWARE: Microsoft Word 97

;; SEQ ID NO 32

;; LENGTH: 504

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-09-306-828-32

## Query Match

2.1%; Score 7; DB 4; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 NLRDKS 266  
|||  
Db 144 NLRDKS 150

## RESULT 46

US-09-489-039A-12958  
; Sequence 12958, Application US/09489039A  
; Patent No. 6610836

## GENERAL INFORMATION:

;; APPLICANT: Gary Breton et. al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
;; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: 2709.2004001

;; CURRENT APPLICATION NUMBER: US/09/489,039A

;; CURRENT FILING DATE: 2000-01-27

;; PRIOR APPLICATION NUMBER: US 60/117,747

;; PRIOR FILING DATE: 1999-01-29

;; NUMBER OF SEQ ID NOS: 14342

;; SEQ ID NO 12958

;; LENGTH: 510

;; TYPE: PRT

;; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12958

## Query Match

2.1%; Score 7; DB 4; Length 510;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 APQIALR 141  
|||  
Db 329 APQIALR 335

## RESULT 47

US-09-800-729-124  
; Sequence 124, Application US/09800729  
; Patent No. 6605532

## GENERAL INFORMATION:

; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 124  
; LENGTH: 514  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-124

Query Match 2.1%; Score 7; DB 4; Length 514;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQE 72  
DB 392 VAQLAQE 398

RESULT 48  
US-09-741-243C-2  
; Sequence 2, Application US/09741243C  
; Patent No. 6399352  
; GENERAL INFORMATION:  
; APPLICANT: Crawford Jr., John Milton  
; APPLICANT: Rice, John  
; APPLICANT: Sevala, Veeresh  
; APPLICANT: Stewart, Sandy  
; TITLE OF INVENTION: NOVEL PLANT PORPHOBILINOGEN AND FUSION  
; FILE REFERENCE: 2022US  
; CURRENT APPLICATION NUMBER: US/09/741,243C  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 551  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: A plant thioredoxin-porphobilinogen  
; OTHER INFORMATION: synthase fusion protein  
US-09-741-243C-2

Query Match 2.1%; Score 7; DB 4; Length 551;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LLRDKSP 267  
DB 483 LLRDKSP 489

RESULT 49  
US-09-489-039A-12036  
; Sequence 12036, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12036  
; LENGTH: 620  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12036

Query Match 2.1%; Score 7; DB 4; Length 620;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 QSLKLLG 233  
DB 71 QSLKLLG 77

RESULT 50  
US-08-026-138E-3  
; Sequence 3, Application US/08026138E  
; Patent No. 5502166  
; GENERAL INFORMATION:  
; APPLICANT: Masayoshi MISHINA  
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nishiohata Residence 1-107  
; STREET: 5214, Nishiohata-machi  
; CITY: Niigata-shi  
; STATE: Niigata-ken  
; COUNTRY: JAPAN  
; ZIP: 951  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS v.5  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/026,138E  
; FILING DATE: 26-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 39563/1992  
; FILING DATE: 26-FEB-1992  
; APPLICATION NUMBER: JP 173155/1992  
; FILING DATE: 30-JUN-1992  
; APPLICATION NUMBER: JP 215017/1992  
; FILING DATE: 12-AUG-1992  
; APPLICATION NUMBER: JP 303878/1992  
; FILING DATE: 13-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hamburg, C. Bruce  
; REGISTRATION NUMBER: 22,389  
; REFERENCE/DOCKET NUMBER: F-4551  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 986-2340  
; TELEFAX: (212) 953-7733  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1239 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single strand  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: mouse  
; TISSUE TYPE: brain  
; PUBLICATION INFORMATION:  
; AUTHORS: Masayoshi MISHINA  
; TITLE: NOVEL PROTEINS AND GENES CODING THE SAME  
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 to 1239  
US-08-026-138E-3

Query Match 2.1%; Score 7; DB 1; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TEAVAL 69  
DB 101 TEAVAL 107

## RESULT 51

US-09-800-729-89  
; Sequence 89, Application US/09800729  
; Patent No. 6605592

## ; GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: 32 Human secreted proteins

; FILE REFERENCE: P2044P1

; CURRENT APPLICATION NUMBER: US/09/800,729

; CURRENT FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: PCT/US00/26013

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: 60/155,709

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 217

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 89

; LENGTH: 1745

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-800-729-89

Query Match 2.1%; Score 7; DB 4; Length 1745;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 72  
DB 1084 VAQLAQ 1090

## RESULT 52

US-09-489-039A-14067

; Sequence 14067, Application US/09489039A

; Patent No. 6610836

## ; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 14067

; LENGTH: 3168

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-14067

Query Match 2.1%; Score 7; DB 4; Length 3168;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 NLALIK 30  
DB 1959 NLALIK 1965

## RESULT 53

US-09-091-501B-10

; Sequence 10, Application US/09091501B

; Patent No. 6518413

; GENERAL INFORMATION:  
; APPLICANT: Tinsley, Jonathon M  
; APPLICANT: Davies, Kay E  
; TITLE OF INVENTION: Utrophin gene expression

; FILE REFERENCE: 620-42

; CURRENT APPLICATION NUMBER: US/09/091,501B

; CURRENT FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: PCT/GB96/03156

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: GB 9525962.8

; PRIOR FILING DATE: 1995-12-19

; PRIOR APPLICATION NUMBER: GB 9615797.9

; PRIOR FILING DATE: 1996-07-26

; PRIOR APPLICATION NUMBER: GB 9622174.2

; PRIOR FILING DATE: 1996-10-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10

; LENGTH: 3433

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: [239] ... [250]

; OTHER INFORMATION: Description of Artificial Sequence: Full length

; OTHER INFORMATION: utrophin construct; Xaa = unknown

US-09-091-501B-10

Query Match 2.1%; Score 7; DB 4; Length 3433;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKDN 24  
DB 1182 VKILKDN 1188

## RESULT 54

US-08-432-871C-75

; Sequence 75, Application US/08432871C

; Patent No. 5877010

## ; GENERAL INFORMATION:

; APPLICANT: Loeb, Lawrence A.

; APPLICANT: Black, Margaret E.

; TITLE OF INVENTION: THYMIDINE KINASE MUTANTS

; NUMBER OF SEQUENCES: 104

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed and Berry LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: US

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/432,871C

; FILING DATE: 02-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: McMasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 240052.409C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; TELEX: 3723836

; INFORMATION FOR SEQ ID NO: 75:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15 amino acids

;  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-432-871C-75

Query Match 1.8%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 LILDRH 240  
Db 1 LILDRH 6

## RESULT 55

US-08-432-871C-82  
; Sequence 82, Application US/08432871C  
; Patent No. 5877010  
; GENERAL INFORMATION:  
; APPLICANT: Loeb, Margaret E.  
; APPLICANT: Black, Margaret E.  
; TITLE OF INVENTION: THYMIDINE KINASE MUTANTS  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/432,871C  
; FILING DATE: 02-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 240052.409C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-432-871C-82

Query Match 1.8%; Score 6; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 LILDRH 240  
Db 1 LILDRH 6

## RESULT 56

US-09-270-956-75  
; Sequence 75, Application US/09270956  
; Patent No. 6451571  
; GENERAL INFORMATION:  
; APPLICANT: Loeb, Margaret E.  
; APPLICANT: Black, Margaret E.  
; TITLE OF INVENTION: THYMIDINE KINASE MUTANTS  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP

;  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/270,956  
; FILING DATE: 17-MAR-1999  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 240052.409C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836  
; INFORMATION FOR SEQ ID NO: 75:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-270-956-75

Query Match 1.8%; Score 6; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 LILDRH 240  
Db 1 LILDRH 6

## RESULT 57

US-09-270-956-82  
; Sequence 82, Application US/09270956  
; Patent No. 6451571  
; GENERAL INFORMATION:  
; APPLICANT: Loeb, Margaret E.  
; APPLICANT: Black, Margaret E.  
; TITLE OF INVENTION: THYMIDINE KINASE MUTANTS  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/270,956  
; FILING DATE: 17-MAR-1999  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McMasters, David D.  
; REGISTRATION NUMBER: 33,963  
; REFERENCE/DOCKET NUMBER: 240052.409C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-270-956-82

Query Match
Best Local Similarity 1.8%; Score 6; DB 4; Length 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 LILDRH 240
Db 1 LILDRH 6

RESULT 58
US-07-965-663A-4
; Sequence 4, Application US/07965663A
; Patent No. 5424290
; GENERAL INFORMATION:
; APPLICANT: Lee, Maloy W.
; TITLE OF INVENTION: No. 5424290el Biologically Active Peptides and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: 1300 I Street, N.W.
; STREET: Washington
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,663A
; FILING DATE: 26-OCT-1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0039-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: /note="May be a C-terminal amide, and/or may
US-07-965-663A-4

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILK 22
Db 16 IVKILK 21

RESULT 59
US-08-033-873-8
; Sequence 8, Application US/08033873
; Patent No. 5459235
; GENERAL INFORMATION:
; APPLICANT: SELSTED, MICHAEL E.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/033,873
; FILING DATE: 19930319
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 9552
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-033-873-8

Query Match
Best Local Similarity 1.8%; Score 6; DB 1; Length 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RROIGT 111
Db 23 RROIGT 28

RESULT 60
US-08-356-832-8
; Sequence 8, Application US/08356832
; Patent No. 5821224
; GENERAL INFORMATION:
; APPLICANT: SELSTED, MICHAEL E.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,832
; FILING DATE: 12-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,873
```



;; FILING DATE: 19-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: CAMPBELL, CATHRYN A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-UC 9552  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 38 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-356-832-8

Query Match 1.8%; Score 6; DB 2; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111  
DB 23 RQIGT 28

RESULT 61  
US-08-988-705-8  
; Sequence 8, Application US/08988705  
; Patent No. 6211148  
; GENERAL INFORMATION:  
; APPLICANT: Selsted, Michael E.  
; APPLICANT: Cullor, James S.  
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL & FLORES, LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/988,705  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/033,873  
; FILING DATE: 19-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/356,832  
; FILING DATE: 13-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UC 2918  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)535-9001  
; TELEFAX: (619)535-8949  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-988-705-8

Query Match 1.8%; Score 6; DB 3; Length 38;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111  
DB 23 RQIGT 28

RESULT 62  
US-08-033-873-7  
; Sequence 7, Application US/08033873  
; Patent No. 5459235  
; GENERAL INFORMATION:  
; APPLICANT: SELSTED, MICHAEL E.  
; APPLICANT: CULLOR, JAMES S.  
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/033,873  
; FILING DATE: 19930319  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UC 9552  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-033-873-7

Query Match 1.8%; Score 6; DB 1; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111  
DB 25 RQIGT 30

RESULT 63  
US-08-033-873-9  
; Sequence 9, Application US/08033873  
; Patent No. 5459235  
; GENERAL INFORMATION:  
; APPLICANT: SELSTED, MICHAEL E.  
; APPLICANT: CULLOR, JAMES S.  
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA

```
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/033,873
; FILING DATE: 19930319
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 9552
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-033-873-9

Query Match 1.8%; Score 6; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 RQIGT 111
Db 25 RQIGT 30

RESULT 64
US-08-356-832-7
; Sequence 7, Application US/08356832
; Patent No. 5821224
; GENERAL INFORMATION:
; APPLICANT: SELSTED, MICHAEL E.
; APPLICANT: CULLOR, JAMES S.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,832
; FILING DATE: 12-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,873
; FILING DATE: 19-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 9552
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-356-832-9

Query Match 1.8%; Score 6; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 RQIGT 111
Db 25 RQIGT 30

RESULT 66
US-08-988-705-7
; Sequence 7, Application US/08988705
; Patent No. 6211148
; GENERAL INFORMATION:
```

```
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-356-832-7

Query Match 1.8%; Score 6; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 RQIGT 111
Db 25 RQIGT 30

RESULT 65
US-08-356-832-9
; Sequence 9, Application US/08356832
; Patent No. 5821224
; GENERAL INFORMATION:
; APPLICANT: SELSTED, MICHAEL E.
; APPLICANT: CULLOR, JAMES S.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,832
; FILING DATE: 12-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,873
; FILING DATE: 19-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 9552
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-356-832-9

Query Match 1.8%; Score 6; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 RQIGT 111
Db 25 RQIGT 30

RESULT 66
US-08-988-705-7
; Sequence 7, Application US/08988705
; Patent No. 6211148
; GENERAL INFORMATION:
```

```
/
/ APPLICANT: Selsted, Michael E.
/ APPLICANT: Cullor, James S.
/ TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
/ TITLE OF INVENTION: NEUTROPHILS
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CAMPBELL & FLORES, LLP
/ STREET: 4370 La Jolla Village Drive, Suite 700
/ CITY: San Diego
/ STATE: California
/ COUNTRY: United States
/ ZIP: 92122
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US 08/988,705
/ FILING DATE: 19-MAR-1993
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/033,873
/ FILING DATE: 13-DEC-1994
/ PRIORITY/DOCKET NUMBER: P-UC 2918
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/DOCKET NUMBER: P-UC 2918
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619)535-8949
/ TELEFAX: (619)535-8949
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-988-705-7

Query Match 1.8%; Score 6; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111
Db 25 RQIGT 30

RESULT 67
US-08-988-705-9
/ Sequence 9, Application US/08988705
/ Patent No. 6211148
/ GENERAL INFORMATION:
/ APPLICANT: Selsted, Michael E.
/ APPLICANT: Cullor, James S.
/ TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES FROM BOVINE
/ TITLE OF INVENTION: NEUTROPHILS
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CAMPBELL & FLORES, LLP
/ STREET: 4370 La Jolla Village Drive, Suite 700
/ CITY: San Diego
/ STATE: California
/ COUNTRY: United States
/ ZIP: 92122
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/988,705
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/033,873
/ FILING DATE: 19-MAR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/356,832
/ FILING DATE: 13-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Campbell, Cathryn A.
/ REGISTRATION NUMBER: 31,815
/ REFERENCE/DOCKET NUMBER: P-UC 2918
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619)535-9001
/ TELEFAX: (619)535-8949
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-988-705-9

Query Match 1.8%; Score 6; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RQIGT 111
Db 25 RQIGT 30

RESULT 68
US-09-057-762-21
/ Sequence 21, Application US/09057762
/ Patent No. 5879909
/ GENERAL INFORMATION:
/ APPLICANT: PERL, ANDRAS
/ TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH
/ TITLE OF INVENTION: A FUNCTION IN METABOLISM
/ NUMBER OF SEQUENCES: 24
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORRISON & FOERSTER
/ STREET: 2000 Pennsylvania Avenue N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20006-1812
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/057,762
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/326,119
/ FILING DATE: 19-OCT-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: LIVNAT, SHMUEL
/ REGISTRATION NUMBER: 33,949
/ REFERENCE/DOCKET NUMBER: 280932000100
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 887-1500
/ TELEFAX: (202) 822-0168
/ TELEX: 90-4030
/ INFORMATION FOR SEQ ID NO: 21:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46 amino acids
```

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-057-762-21

Query Match 1.8%; Score 6; DB 2; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLLGEL 235  
DB 15 KLLGEL 20

## RESULT 69

US-08-326-119A-21  
; Sequence 21, Application US/08326119A  
; Patent No. 6018021  
; GENERAL INFORMATION:  
; APPLICANT: PERL, ANDRAS  
; TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH  
; TITLE OF INVENTION: A FUNCTION IN METABOLISM  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/326,119A  
; FILING DATE: 19-OCT-1994  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: LIVNAT, SHMUEL  
; REGISTRATION NUMBER: 33,949  
; REFERENCE/DOCKET NUMBER: 280932000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 822-0168  
; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 46 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-326-119A-21

Query Match 1.8%; Score 6; DB 3; Length 46;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLLGEL 235  
DB 15 KLLGEL 20

## RESULT 70

US-09-621-976-6437  
; Sequence 6437, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6437  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 30  
; OTHER INFORMATION: Xaa = Ala,Gly  
; NAME/KEY: UNSURE  
; LOCATION: 38  
; OTHER INFORMATION: Xaa = Phe,Val  
US-09-621-976-6437

Query Match 1.8%; Score 6; DB 4; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
DB 15 SLKLLG 20

## RESULT 71

US-09-621-976-6438  
; Sequence 6438, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6438  
; LENGTH: 52  
; TYPE: PRT

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 30  
; OTHER INFORMATION: Xaa = Ala,Gly  
; NAME/KEY: UNSURE  
; LOCATION: 51  
; OTHER INFORMATION: Xaa = Gly,Val  
US-09-621-976-6438

Query Match 1.8%; Score 6; DB 4; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
DB 15 SLKLLG 20

## RESULT 72

US-09-543-681A-6353  
; Sequence 6353, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6353  
; LENGTH: 60  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6353

Query Match 1.8%; Score 6; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 IILFSN 163  
| | | | |  
Db 17 IILFSN 22

RESULT 73  
US-09-107-532A-5451  
; Sequence 5451, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5451:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES

ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:

NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...61  
SEQUENCE DESCRIPTION: SEQ ID NO: 5451:  
US-09-107-532A-5451

Query Match 1.8%; Score 6; DB 4; Length 61;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LEKODK 33  
| | | | |  
Db 16 LEKODK 21

RESULT 74  
US-09-673-395A-510  
; Sequence 510, Application US/09673395A  
; Patent No. 6620923  
; GENERAL INFORMATION:  
; APPLICANT: SPECHT, THOMAS  
; APPLICANT: HINZMANN, BERND  
; APPLICANT: SCHMITT, ARMIN  
; APPLICANT: PILARSKY, CHRISTIAN  
; APPLICANT: DAHL, EDGAR  
; APPLICANT: ROSENTHAL, ANDRE  
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE  
; FILE REFERENCE: ALBRE-12  
; CURRENT APPLICATION NUMBER: US/09/673,395A  
; CURRENT FILING DATE: 2000-10-17  
; NUMBER OF SEQ ID NOS: 637  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 510  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-673-395A-510

Query Match 1.8%; Score 6; DB 4; Length 63;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
| | | | |  
Db 1 SLKLLG 6

RESULT 75  
US-09-091-814-25  
; Sequence 25, Application US/09091814  
; Patent No. 6218513  
; GENERAL INFORMATION:  
; APPLICANT: Anthony-Cahill, Spencer J.  
; APPLICANT: Epp, Janet K  
; APPLICANT: Kerwin, Bruce A.  
; APPLICANT: Olin O., Peter  
; APPLICANT: Mathews J., Antony  
; TITLE OF INVENTION: GLOBINS CONTAINING BINDING DOMAINS  
; FILE REFERENCE: BXTB2005  
; CURRENT APPLICATION NUMBER: US/09/091,814  
; CURRENT FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 64  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GCN4 no Cys  
US-09-091-814-25

Query Match 1.8%; Score 6; DB 3; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.le+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LKLIGE 234  
| | | | |  
Db 41 LKLIGE 46

RESULT 76  
US-09-134-001C-4534

```
; Sequence 4534, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4534
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4534

Query Match
; Sequence 66, Score 6; DB 4; Length 66;
; Best Local Similarity 100.0%; Pred. No. 2.1e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82
DB 14 GLLVTL 19

RESULT 77
US-09-621-976-5578
; Sequence 5578, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5578
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -14...-1
US-09-621-976-5578

Query Match
; Sequence 66, Score 6; DB 4; Length 73;
; Best Local Similarity 100.0%; Pred. No. 2.3e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 PTEAVA 67
DB 58 PTEAVA 63

RESULT 78
US-09-858-664A-32
; Sequence 32, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
```

```
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-32

Query Match
; Sequence 66, Score 6; DB 4; Length 82;
; Best Local Similarity 100.0%; Pred. No. 2.6e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKN 297
DB 64 EILLKN 69

RESULT 79
US-10-274-978-33
; Sequence 33, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Human
US-10-274-978-33

Query Match
; Sequence 66, Score 6; DB 4; Length 82;
; Best Local Similarity 100.0%; Pred. No. 2.6e+02;
; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 292 EILLKN 297
DB 64 EILLKN 69

RESULT 80
US-09-134-000C-6127
; Sequence 6127, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6127
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6127
```

Query Match 1.8%; Score 6; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 154 PLAKII 159  
Db 58 PLAKII 63

RESULT 81  
US-09-732-210-888  
; Sequence 888, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Sealle, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 888  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Methanococcus jannaschii  
US-09-732-210-888

Query Match 1.8%; Score 6; DB 4; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 15 ABEIKI 20  
Db 45 ABEIKI 50

RESULT 82  
US-09-107-532A-4435  
; Sequence 4435, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS: 7310  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 4435:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 93 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...93  
SEQUENCE DESCRIPTION: SEQ ID NO: 4435:  
US-09-107-532A-4435

Query Match 1.8%; Score 6; DB 4; Length 93;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 5 PLFSKS 10  
Db 37 PLFSKS 42

RESULT 83  
US-09-621-976-4269  
; Sequence 4269, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4269  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-4269

Query Match 1.8%; Score 6; DB 4; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 322 KNYLIK 327  
Db 20 KNYLIK 25

RESULT 84  
US-09-621-976-5842  
; Sequence 5842, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm

; SEQ ID NO 5842  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -51...-1  
US-09-621-976-5842

Query Match 1.8%; Score 6; DB 4; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLLTRH 194  
DB 54 DLLTRH 59

## RESULT 85

US-09-252-991A-18411  
; Sequence 18411, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18411

; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18411

Query Match 1.8%; Score 6; DB 4; Length 114;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 EXLLOS 219  
DB 31 EXLLOS 36

## RESULT 86

US-08-961-083-68  
; Sequence 68, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:

; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 117 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-961-083-68

Query Match 1.8%; Score 6; DB 3; Length 117;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 FEGKQD 96  
DB 19 FEGKQD 24

## RESULT 87

US-08-936-165A-347  
; Sequence 347, Application US/08936165A  
; Patent No. 6348582  
; GENERAL INFORMATION:

; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Lonetto, Michael  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Richard  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Ward, Judith  
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,  
; TITLE OF INVENTION: Polypeptides and Their Uses  
; NUMBER OF SEQUENCES: 534  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-Seq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/936,165A  
; FILING DATE: 24-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/027,032  
; FILING DATE: 24-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gilmi, Edward R  
; REGISTRATION NUMBER: 38,891  
; REFERENCE/DOCKET NUMBER: P50549  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-4478  
; TELEFAX: 610-270-5090  
; TELEX:



; INFORMATION FOR SEQ ID NO: 347:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 117 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Protein

US-08-936-165A-347

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 117;

Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 18 VKILKD 23

Db 1 VKILKD 6

RESULT 88

US-09-536-784-68

; Sequence 68, Application US/09536784

; Patent No. 6573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 68:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 117 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 68:

US-09-536-784-68

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 117;

Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 91 FECKKD 96

Db 19 FECKKD 24

RESULT 89

US-08-513-974B-50

; Sequence 50, Application US/08513974B

; Patent No. 6114139

; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji

; APPLICANT: Hosoya, Masaki

; APPLICANT: Fujii, Ryo

; APPLICANT: Ohtaki, Tetsuya

; APPLICANT: Fukusumi, Shoji

; APPLICANT: Ohgi, Kazuhiro

; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,

; NUMBER OF SEQUENCES: 380

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/513,974B

; FILING DATE: 14-SEP-1995

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP95/01599

; FILING DATE: 10-AUG-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-093989

; FILING DATE: 19-AUG-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-057186

; FILING DATE: 16-MAR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-007177

; FILING DATE: 20-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-326611

; FILING DATE: 28-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-270017

; FILING DATE: 02-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-236357

; FILING DATE: 30-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-236356

; FILING DATE: 30-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-189274

; FILING DATE: 11-AUG-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-189273

; FILING DATE: 11-AUG-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-189272

; FILING DATE: 11-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Resnick, David S.

; REGISTRATION NUMBER: 34,235

; REFERENCE/DOCKET NUMBER: 45753

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 119 amino acids

; TYPE: amino acid

; STRANDEDNESS:

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-513-974B-50

Query Match 1.8%; Score 6; DB 3; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 KDVTQI 100
Db 45 KDVTQI 50

RESULT 90
US-08-513-974B-354
; Sequence 354, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 354:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-513-974B-354

Query Match 1.8%; Score 6; DB 3; Length 119;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 KDVTQI 100
Db 45 KDVTQI 50

RESULT 91
US-08-513-974B-355
; Sequence 355, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
```

APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1945  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Reenick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: peptide  
US-08-513-974B-355

Query Match 1.8%; Score 6; DB 3; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 KDVTQI 100  
Db 45 KDVTQI 50

RESULT 92  
US-09-461-436B-50  
Sequence 50, Application US/09461436B  
Patent No. 6538107  
GENERAL INFORMATION:  
APPLICANT: Shuji Hinuma  
Yasuaki Ito  
Ryo Fujii  
TITLE OF INVENTION: G Protein Coupled Receptor Protein,  
Production, And Use thereof  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edwards & Angell, LLP  
STREET: 101 Federal Street  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02209  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/461.436B  
FILING DATE: 14-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/513,974  
FILING DATE: 14-SEP-1995  
APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995  
APPLICATION NUMBER: 7-093989  
FILING DATE: 19-APR-1995  
APPLICATION NUMBER: 7-057186  
FILING DATE: 16-MAR-1995  
APPLICATION NUMBER: 7-007177  
FILING DATE: 20-JAN-1995  
APPLICATION NUMBER: 6-326611  
FILING DATE: 28-DEC-1994  
APPLICATION NUMBER: 6-270017  
FILING DATE: 02-NOV-1994  
APPLICATION NUMBER: 6-236357  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-236356  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-189274  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189273  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CONLIN, DAVID G.  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 45753 DIV2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-439-4444  
TELEFAX: 617-439-4170  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-09-461-436B-50

Query Match 1.8%; Score 6; DB 4; Length 119;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 KDVTQI 100  
Db 45 KDVTQI 50

RESULT 93  
US-09-134-000C-3508  
Sequence 3508, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 3508  
LENGTH: 128  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-3508

Query Match 1.8%; Score 6; DB 4; Length 128;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 LLKNQP 299

```
Db      86 LLKNQP 91
|||||
RESULT 94
US-09-134-000C-5393
; Sequence 5393, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5393
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5393
Query Match      1.8%; Score 6; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      105 LRRQIG 110
Db      86 LRRQIG 91
|||||
RESULT 95
US-09-481-161-4
; Sequence 4, Application US/09481161A
; Patent No. 6337432
; GENERAL INFORMATION:
; APPLICANT: Dahlman, Douglas
; APPLICANT: Webb, Bruce
; APPLICANT: Maiti, Indu
; TITLE OF INVENTION: Materials and Methods Useful to Affect Growth and Development of
; FILE REFERENCE: P-1023
; CURRENT APPLICATION NUMBER: US/09/481,161A
; CURRENT FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Microplitis sp.
US-09-481-161-4
Query Match      1.8%; Score 6; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      118 YISAHP 123
Db      19 YISAHP 24
|||||
RESULT 96
US-09-328-352-4824
; Sequence 4824, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NAEMANNII ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4824
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4824
Query Match      1.8%; Score 6; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      156 AKIILF 161
Db      79 AKIILF 84
|||||
RESULT 97
US-09-634-238-372
; Sequence 372, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043UI
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 372
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-372
Query Match      1.8%; Score 6; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      248 YISKPE 253
Db      61 YISKPE 66
|||||
RESULT 98
US-09-198-452A-673
; Sequence 673, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; FILE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 673
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
```

; LOCATION: 1...133  
; OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-673

Query Match 1.8%; Score 6; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILK 22  
Db 37 IVKILK 42

RESULT 99  
US-09-732-210-804  
; Sequence 804, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 804  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Pyrobotrys stellata  
US-09-732-210-804

Query Match 1.8%; Score 6; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 VTKRQS 228  
Db 53 VTKRQS 58

RESULT 100  
US-09-732-210-1265  
; Sequence 1265, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 1265  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Euglena gracilis  
US-09-732-210-1265

Query Match 1.8%; Score 6; DB 4; Length 134;

Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AKILF 161  
Db 17 AKILF 22

RESULT 101  
US-08-383-621-5  
; Sequence 5, Application US/08383621  
; Patent No. 5951972  
; GENERAL INFORMATION:  
; APPLICANT: Daley, Michael J.  
; APPLICANT: Buckwalter, Brian L.  
; APPLICANT: Cady, Susan M.  
; APPLICANT: Shieh, Hong-Ming  
; APPLICANT: Bohlen, Peter  
; APPLICANT: Seddon, Andrew P.  
; TITLE OF INVENTION: Stabilization Of Somatotropins And Other  
; TITLE OF INVENTION: Proteins By Modification Of Cysteine Residues  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. Estelle J. Tsevdos  
; STREET: 1937 West Main Street, P.O. Box 60  
; CITY: Stamford  
; STATE: Connecticut  
; COUNTRY: U.S.A.  
; ZIP: 06904-0060  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/383,621  
; FILING DATE: 06-FEB-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER: US 07/766,142  
; FILING DATE: 25-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Tsevdos, Estelle J.  
; REGISTRATION NUMBER: 31,145  
; REFERENCE/DOCKET NUMBER: 31,278-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 203-321-2756  
; TELEFAX: 203-321-2971  
; TELEX: 203-710-474-4059  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 135 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-383-621-5

Query Match 1.8%; Score 6; DB 2; Length 135;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLKL 257  
Db 31 PENLKL 36

RESULT 102  
US-08-459-906-5  
; Sequence 5, Application US/08459906  
; Patent No. 6010999  
; GENERAL INFORMATION:  
; APPLICANT: Daley, Michael J.  
; APPLICANT: Buckwalter, Brian L.

APPLICANT: Cady, Susan M.  
APPLICANT: Shieh, Hong-Ming  
APPLICANT: Bohnen, Peter  
APPLICANT: Seddon, Andrew P.  
TITLE OF INVENTION: Stabilization of Somatotropins and Other  
TITLE OF INVENTION: Proteins by Modification of Cysteine Residues  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,906  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Darryl L.  
REGISTRATION NUMBER: 34,276  
REFERENCE/DOCKET NUMBER: 31,278-03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-831-3247  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 135 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-906-5

Query Match 1.8%; Score 6; DB 3; Length 135;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 252 PENLKL 257  
Db 31 PENLKL 36  
|||||

RESULT 103  
US-09-149-476-554  
Sequence 554, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PC/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877

EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 1.8%; Score 6; DB 4; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TLIADL 86  
DB 100 TLIADL 105

## RESULT 104

US-09-489-039A-11504  
; Sequence 11504, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11504  
; LENGTH: 136  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11504

Query Match 1.8%; Score 6; DB 4; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKILG 233  
DB 102 SLKILG 107

## RESULT 105

US-09-489-039A-13120  
; Sequence 13120, Application US/09489039A  
; Patent No. 6610836

```
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13120
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-039A-13120

Query Match 1.8%; Score 6; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 VEYISA 121
DB 91 VEYISA 96

RESULT 106
US-09-370-838-73
; Sequence 73, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Sacrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-73

Query Match 1.8%; Score 6; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 EDYEKL 216
DB 58 EDYEKL 63

RESULT 107
US-08-872-783-1
; Sequence 1, Application US/08872783
; Patent No. 5858717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN FORMIN BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
```

```
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,783
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0317 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TESTTUT02
; CLONE: 2345085
US-08-872-783-1

Query Match 1.8%; Score 6; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 DLKXTA 336
DB 92 DLKXTA 97

RESULT 108
US-08-858-207A-423
; Sequence 423, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
```



```
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 423:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-423

Query Match      1.8%; Score 6; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      195 KVLVAD 200
DB      53 KVLVAD 58
      |||||

RESULT 109
US-09-252-991A-30905
; Sequence 30905, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30905
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30905

Query Match      1.8%; Score 6; DB 4; Length 155;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      138 IALRCG 143
DB      118 IALRCG 123
      |||||

RESULT 110
US-09-134-001C-4854
; Sequence 4854, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4854
```

```
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4854

Query Match      1.8%; Score 6; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      78 LLVTLL 83
DB      16 LLVTLL 21
      |||||

RESULT 111
US-09-732-210-532
; Sequence 532, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 532
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Mycoplasma pneumoniae
US-09-732-210-532

Query Match      1.8%; Score 6; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      102 NNILR 107
DB      56 NNILR 61
      |||||

RESULT 112
US-09-732-210-531
; Sequence 531, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 531
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Mycoplasma genitalium
US-09-732-210-531
```

Query Match 1.8%; Score 6; DB 4; Length 162;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 NNILRR 107  
      |||||  
Db 57 NNILRR 62

## RESULT 113

US-09-328-352-5779  
; Sequence 5779, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: GTC99-03PA  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 3252  
; SEQ ID NO 5779  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5779

Query Match 1.8%; Score 6; DB 4; Length 163;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVSKSL 46  
      |||||  
Db 66 EVSKSL 71

## RESULT 114

US-09-252-991A-19902  
; Sequence 19902, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19902  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19902

Query Match 1.8%; Score 6; DB 4; Length 170;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 FKDLTL 192  
      |||||  
Db 95 FKDLTL 100

## RESULT 115

US-09-149-476-526  
; Sequence 526, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: PZ002PA  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 1.8%; Score 6; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 KVLVAD 200  
Db 40 KVLVAD 45  
|||||

## RESULT 116

US-09-328-352-4965  
; Sequence 4965, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4965

```
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4965

Query Match      1.8%; Score 6; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23
Db 83 VKILKD 88

RESULT 117
US-09-107-532A-5800
; Sequence 5800, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Atinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5800:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...178
; SEQUENCE DESCRIPTION: SEQ ID NO: 5800:
US-09-107-532A-5800

Query Match      1.8%; Score 6; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 KTQPIV 291
Db 100 KTQPIV 105
```

```
RESULT 118
US-09-461-325-196
; Sequence 196, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029PI
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-196
```

```
Query Match      1.8%; Score 6; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 77 GLLVTL 82
Db 99 GLLVTL 104
```

```
RESULT 119
US-10-012-542-196
; Sequence 196, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029PI
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-196
```

Query Match 1.8%; Score 6; DB 4; Length 181;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
|||||  
Db 99 GLLVTL 104  
|||||

## RESULT 120

US-09-198-452A-612  
; Sequence 612, Application US/09198452A

; Patent No. 6559294

; GENERAL INFORMATION:

; APPLICANT: Grifais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; FILE OF INVENTION: and treatment of infection

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 612

; LENGTH: 183

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-198-452A-612

Query Match 1.8%; Score 6; DB 4; Length 183;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 VTLIAD 85  
|||||  
Db 50 VTLIAD 55  
|||||

## RESULT 121

US-09-489-039A-7656

; Sequence 7656, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 7656

; LENGTH: 184

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-7656

Query Match 1.8%; Score 6; DB 4; Length 184;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 DLLTRH 194  
|||||  
Db 47 DLLTRH 52  
|||||

## RESULT 122

US-09-489-039A-12490

; Sequence 12490, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 12490

; LENGTH: 184

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12490

Query Match 1.8%; Score 6; DB 4; Length 184;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 VKILKD 23  
|||||  
Db 86 VKILKD 91  
|||||

## RESULT 123

US-09-489-039A-8986

; Sequence 8986, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8986

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8986

Query Match 1.8%; Score 6; DB 4; Length 188;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 GELILD 238  
|||||  
Db 41 GELILD 46  
|||||

## RESULT 124

US-09-107-532A-4452

; Sequence 4452, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 4452:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 190 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...190  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4452:  
US-09-107-532A-4452

Query Match 1.8%; Score 6; DB 4; Length 190;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LEQDKX 33  
|||||  
Db 40 LEQDKX 45

## RESULT 125

US-09-252-991A-22287  
; Sequence 22287, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22287  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22287

Query Match 1.8%; Score 6; DB 4; Length 207;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LLRDKS 266  
|||||  
Db 125 LLRDKS 130

## RESULT 126

US-08-811-519-24  
; Sequence 24, Application US/08811519B

; Patent No. 6630345  
; GENERAL INFORMATION:  
; APPLICANT: Petrenko, Alexandre  
; TITLE OF INVENTION: CALCIUM INDEPENDENT RECEPTOR OF ALPHA-LATROTOXIN,  
; TITLE OF INVENTION: CHARACTERIZATION AND USES THEREOF  
; FILE REFERENCE: 1049-1-007  
; CURRENT APPLICATION NUMBER: US/08/811,519B  
; CURRENT FILING DATE: 1997-03-04  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-811-519-24

Query Match 1.8%; Score 6; DB 4; Length 207;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 FLSSFQ 309  
|||||  
Db 181 FLSSFQ 186

## RESULT 127

US-09-134-001C-3080  
; Sequence 3080, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3080  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3080

Query Match 1.8%; Score 6; DB 4; Length 209;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 TLIADL 86  
|||||  
Db 176 TLIADL 181

## RESULT 128

US-09-391-741A-10  
; Sequence 10, Application US/09391741A  
; Patent No. 6555732  
; GENERAL INFORMATION:  
; APPLICANT: Duvick, Jonathan P.  
; APPLICANT: Sharma, Yogesh Kumar  
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use  
; FILE REFERENCE: 0866D  
; CURRENT APPLICATION NUMBER: US/09/391,741A  
; CURRENT FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: 60/111,919  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 60/100,284  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: 09/391,741  
; PRIOR FILING DATE: 1999-09-08

```
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Zea mays
US-09-391-741A-10

Query Match          1.8%; Score 6; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110
Db 146 LRRQIG 151

RESULT 129
US-09-391-741A-24
; Sequence 24, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duivick, Jonathan P.
; APPLICANT: Sharma, Yogesh Kumar
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Zea mays
US-09-391-741A-24

Query Match          1.8%; Score 6; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110
Db 146 LRRQIG 151

RESULT 130
US-09-391-741A-34
; Sequence 34, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duivick, Jonathan P.
; APPLICANT: Sharma, Yogesh Kumar
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Zea mays
US-09-391-741A-34

Query Match          1.8%; Score 6; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110
Db 146 LRRQIG 151

RESULT 131
US-09-391-741A-2
; Sequence 2, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duivick, Jonathan P.
; APPLICANT: Sharma, Yogesh Kumar
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
US-09-391-741A-2

Query Match          1.8%; Score 6; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110
Db 148 LRRQIG 153

RESULT 132
US-09-391-741A-16
; Sequence 16, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duivick, Jonathan P.
; APPLICANT: Sharma, Yogesh Kumar
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
US-09-391-741A-16

Query Match          1.8%; Score 6; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110  
|  
|  
|  
|  
|  
Db 148 LRRQIG 153

## RESULT 133

US-09-391-741A-26  
; Sequence 26, Application US/09391741A  
; Patent No. 6555732  
; GENERAL INFORMATION:  
; APPLICANT: Duwick, Jonathan P.  
; APPLICANT: Sharma, Yogesh Kumar  
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use  
; FILE REFERENCE: 0866D  
; CURRENT APPLICATION NUMBER: US/09/391.741A  
; CURRENT FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: 60/111,919  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 60/100,284  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: 09/391.741  
; PRIOR FILING DATE: 1999-09-08  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 214  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-391-741A-26

Query Match 1.8%; Score 6; DB 4; Length 214;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 LRRQIG 110  
|  
|  
|  
|  
|  
Db 148 LRRQIG 153

## RESULT 134

US-09-328-352-6571  
; Sequence 6571, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6571  
; LENGTH: 214  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6571

Query Match 1.8%; Score 6; DB 4; Length 214;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 SGLLVT 81  
|  
|  
|  
|  
|  
Db 116 SGLLVT 121

## RESULT 135

US-09-634-238-341  
; Sequence 341, Application US/09634238  
; Patent No. 6544772  
; GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Iubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christenson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; FILE REFERENCE: 11000.1043U1  
; CURRENT APPLICATION NUMBER: US/09/634,238  
; CURRENT FILING DATE: 2000-08-08  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 341  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Lactobacillus rhamnosus  
US-09-634-238-341

Query Match 1.8%; Score 6; DB 4; Length 218;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 KDLLTR 193  
|  
|  
|  
|  
|  
Db 171 KDLLTR 176

## RESULT 136

US-09-328-352-6222  
; Sequence 6222, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 6222  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-6222

Query Match 1.8%; Score 6; DB 4; Length 218;  
Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 FLEQNY 206  
|  
|  
|  
|  
|  
Db 201 FLEQNY 206

## RESULT 137

US-09-247-373B-50  
; Sequence 50, Application US/09247373B  
; Patent No. 6168954  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; APPLICANT: O'KEEFE, DANIEL  
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CL-1108-A  
; CURRENT APPLICATION NUMBER: US/09/247,373B  
; CURRENT FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 08/924,747  
; PRIOR FILING DATE: 1997-09-05  
; NUMBER OF SEQ ID NOS: 56



```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 50
; LENGTH: 220
; TYPE: PRT
; ORGANISM: SOYBEAN
US-09-247-373B-50

Query Match
  1.8%; Score 6; DB 3; Length 220;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 212 DYEKL 217
    |||||
Db 203 DYEKL 208

RESULT 138
US-09-134-001C-3053
; Sequence 3053, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3053
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3053

Query Match
  1.8%; Score 6; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 329 IRLKK 334
    |||||
Db 88 IRLKK 93

RESULT 139
US-09-489-039A-9910
; Sequence 9910, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9910
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9910

Query Match
  1.8%; Score 6; DB 4; Length 227;
Best Local Similarity 100.0%; Pred. No. 6.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 272 EAFHV 277
    |||||

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 50
; LENGTH: 220
; TYPE: PRT
; ORGANISM: SOYBEAN
US-09-247-373B-50

Query Match
  1.8%; Score 6; DB 3; Length 220;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 212 DYEKL 217
    |||||
Db 203 DYEKL 208

RESULT 140
US-09-107-532A-5238
; Sequence 5238, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...230
; SEQUENCE DESCRIPTION: SEQ ID NO: 5238:
US-09-107-532A-5238

Query Match
  1.8%; Score 6; DB 4; Length 230;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LIRDKS 266
    |||||
Db 214 LIRDKS 219

RESULT 141
US-09-543-681A-6078
; Sequence 6078, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
```

; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6078  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6078

Query Match 1.8%; Score 6; DB 4; Length 231;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LRRQIG 110  
| | | | |  
Db 87 LRRQIG 92

## RESULT 142

US-09-489-039A-10973  
; Sequence 10973, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10973  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10973

Query Match 1.8%; Score 6; DB 4; Length 231;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LRRQIG 110  
| | | | |  
Db 86 LRRQIG 91

## RESULT 143

US-08-468-036-7  
; Sequence 7, Application US/08468036  
; Patent No. 5728806  
; GENERAL INFORMATION:  
; APPLICANT: DeMaggio, Anthony J.  
; APPLICANT: Hoekstra, Merl F.  
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that  
; TITLE OF INVENTION: Interact with Casein Kinase I  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,036  
; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/184,605  
; FILING DATE: 21-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5728806and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/31784  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-468-036-7

Query Match 1.8%; Score 6; DB 1; Length 232;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 KYVELS 175  
| | | | |  
Db 113 KYVELS 118

## RESULT 144

US-08-376-843-7  
; Sequence 7, Application US/08376843  
; Patent No. 5846764  
; GENERAL INFORMATION:  
; APPLICANT: DeMaggio, Anthony J.  
; APPLICANT: Hoekstra, Merl F.  
; TITLE OF INVENTION: Materials and Methods Relating to Proteins  
; TITLE OF INVENTION: that Interact with Casein Kinase I  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/376,843  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/184,605  
; FILING DATE: 21-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5846764and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/31784  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-376-843-7

Query Match 1.8%; Score 6; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 KYVELS 175  
| | | | |  
DB 113 KYVELS 118

RESULT 145  
US-09-134-001C-4013  
; Sequence 4013, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR FILING DATE: 1997-11-08 US 60/064,964  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4013  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4013

Query Match 1.8%; Score 6; DB 4; Length 233;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 NYLIKQ 328  
| | | | |  
DB 124 NYLIKQ 129

RESULT 146  
US-09-198-452A-182  
; Sequence 182, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 182  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 1....242  
; OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-182

Query Match 1.8%; Score 6; DB 4; Length 242;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 IVKILK 22  
| | | | |  
DB 72 IVKILK 77

RESULT 147  
US-09-252-991A-29870  
; Sequence 29870, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29870  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29870

Query Match 1.8%; Score 6; DB 4; Length 243;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLIVTL 82  
| | | | |  
DB 29 GLIVTL 34

RESULT 148  
US-09-100-804-13  
; Sequence 13, Application US/09100804  
; Patent No. 6056472  
; GENERAL INFORMATION:  
; APPLICANT: GONEZ, LEONEL JORGE  
; APPLICANT: SARAS, JAN  
; APPLICANT: CLAESSON-WELSH, LENA  
; APPLICANT: HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
; TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN  
; TITLE OF INVENTION: TYROSINE PHOSPHATASES  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,804  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/596,291  
; FILING DATE: 09-AUG-1996  
; APPLICATION NUMBER: US 08/115,573  
; FILING DATE: 01-SEP-1993  
; PRIOR APPLICATION DATA: PCT/US94/09943  
; APPLICATION NUMBER: 01-SEP-1994  
; FILING DATE: 01-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: L0461/7003  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 247 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-09-100-804-13

Query Match 1.8%; Score 6; DB 3; Length 247;

Best Local Similarity 100.0%; Pred. No. 7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLV 80

Db 169 SSGLLV 174

RESULT 149

US-09-489-039A-7921

; Sequence 7921, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 7921

; LENGTH: 247

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-7921

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 247;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VAQLAQ 71

Db 203 VAQLAQ 208

RESULT 150

US-09-489-039A-8345

; Sequence 8345, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8345

; LENGTH: 249

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-8345

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 249;

Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69

Db 45 EAVAQL 50

RESULT 151

US-09-252-991A-25583

; Sequence 25583, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25583

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-25583

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 254;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 DKASEE 41

Db 189 DKASEE 194

RESULT 152

US-09-252-991A-19444

; Sequence 19444, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 19444

; LENGTH: 255

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-19444

Query Match

Best Local Similarity 1.8%; Score 6; DB 4; Length 255;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLIVTL 82

Db 177 GLIVTL 182

RESULT 153

US-08-311-731A-141

; Sequence 141, Application US/08311731A

/ Patent No. 6583266  
/ GENERAL INFORMATION:  
/ APPLICANT: SMITH, DOUGLAS  
/ APPLICANT: MAO, JEN-I  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
/ TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
/ TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
/ NUMBER OF SEQUENCES: 411  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
/ STREET: 600 ATLANTIC AVENUE  
/ CITY: BOSTON  
/ STATE: MASSACHUSETTS  
/ COUNTRY: USA  
/ ZIP: 02210  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/311,731A  
/ FILING DATE:  
/ CLASSIFICATION: 530  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: GATES, EDWARD R.  
/ REGISTRATION NUMBER: 31,616  
/ REFERENCE/DOCKET NUMBER: C0044/7125  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 617/720-3500  
/ TELEFAX: 617/720-2441  
/ INFORMATION FOR SEQ ID NO: 141:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 255 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ HYPOTHETICAL: YES  
/ ORIGINAL SOURCE:  
/ ORGANISM: Mycobacterium leprae  
/ US-08-311-731A-141

Query Match 1.8%; Score 6; DB 4; Length 255;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LVTLLA 84  
| | | | |  
Db 202 LVTLLA 207

RESULT 154  
US-09-634-955B-23  
/ Sequence 23, Application US/09634955B  
/ Patent No. 6511834  
/ GENERAL INFORMATION:  
/ APPLICANT: Meyers, Rachel  
/ APPLICANT: Cook, William James  
/ TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE  
/ TITLE OF INVENTION: MOLECULES AND USES THEREFOR  
/ FILE REFERENCE: WNI-134  
/ CURRENT APPLICATION NUMBER: US/09/634,955B  
/ CURRENT FILING DATE: 2000-08-08  
/ PRIOR APPLICATION NUMBER: 60/192,002  
/ PRIOR FILING DATE: 2000-03-24  
/ NUMBER OF SEQ ID NOS: 35  
/ SOFTWARE: FastSeq for Windows Version 4.  
/ SEQ ID NO 23  
/ LENGTH: 257  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: 3-beta hydroxysteroid dehydrogenase domain

US-09-634-955B-23

Query Match 1.8%; Score 6; DB 4; Length 257;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 KSPNIQ 270  
| | | | |  
Db 142 KSPNIQ 147

RESULT 155  
US-09-252-991A-25534  
/ Sequence 25534, Application US/09252991A  
/ Patent No. 6551795  
/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ CURRENT FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ PRIOR FILING DATE: 1998-07-27  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 25534  
/ LENGTH: 257  
/ TYPE: PRT  
/ ORGANISM: Pseudomonas aeruginosa  
/ US-09-252-991A-25534

Query Match 1.8%; Score 6; DB 4; Length 257;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLVTLI 83  
| | | | |  
Db 237 LLVTLI 242

RESULT 156  
US-09-328-352-4853  
/ Sequence 4853, Application US/09328352  
/ Patent No. 6562958  
/ GENERAL INFORMATION:  
/ APPLICANT: Gary L. Breton et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
/ TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
/ FILE REFERENCE: GTC99-03PA  
/ CURRENT APPLICATION NUMBER: US/09/328,352  
/ CURRENT FILING DATE: 1999-06-04  
/ NUMBER OF SEQ ID NOS: 8252  
/ SEQ ID NO 4853  
/ LENGTH: 263  
/ TYPE: PRT  
/ ORGANISM: Acinetobacter baumannii  
/ US-09-328-352-4853

Query Match 1.8%; Score 6; DB 4; Length 263;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 QLAQEL 73  
| | | | |  
Db 60 QLAQEL 65

RESULT 157  
US-09-198-452A-23  
/ Sequence 23, Application US/09198452A  
/ Patent No. 6559294

```

; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 23
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-23

Query Match 1.8%; Score 6; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 DYKLL 217
Db 135 DYKLL 140

RESULT 158
US-07-857-224B-42
; Sequence 42, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION:
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; FEATURE: Protein kinase; Table 8 Column 48
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988

US-07-857-224B-42
Query Match 1.8%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KDLAI 27
Db 69 KDLAI 74

RESULT 159
US-09-252-991A-31810
; Sequence 31810, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31810
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31810

Query Match 1.8%; Score 6; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 LVADFL 202
Db 102 LVADFL 107

RESULT 160
US-09-101-146-62
; Sequence 62, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
```

; REFERENCE/DOCKET NUMBER: DC-0050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (856) 810-1515  
; TELEFAX: (856) 810-1454  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-09-101-146-62

Query Match 1.8%; Score 6; DB 3; Length 270;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 YISKPE 253  
Db 192 YISKPE 197

## RESULT 161

US-09-134-001C-2864  
; Sequence 2864, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 2864  
; LENGTH: 272  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2864

Query Match 1.8%; Score 6; DB 4; Length 272;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 TLIADL 86  
Db 72 TLIADL 77

## RESULT 162

US-09-489-039A-13748  
; Sequence 13748, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13748  
; LENGTH: 272  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13748

Query Match 1.8%; Score 6; DB 4; Length 272;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 81 TLIADL 86  
Db 86 TLIADL 91

## RESULT 163

US-09-134-001C-3921  
; Sequence 3921, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3921  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3921

Query Match 1.8%; Score 6; DB 4; Length 274;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 EDYEKL 216  
Db 235 EDYEKL 240

## RESULT 164

US-09-328-352-4229  
; Sequence 4229, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4229  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-4229

Query Match 1.8%; Score 6; DB 4; Length 279;  
Best Local Similarity 100.0%; Pred. No. 7.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 ADFLEQ 204  
Db 153 ADFLEQ 158

## RESULT 165

US-09-328-352-7494  
; Sequence 7494, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7494  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7494

Query Match 1.8%; Score 6; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DNLAIL 28  
Db 135 DNLAIL 140

RESULT 166  
US-09-540-236-3538  
; Sequence 3538, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 3538  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: M.catarrhalis  
US-09-540-236-3538

Query Match 1.8%; Score 6; DB 4; Length 282;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 TDDEQF 318  
Db 59 TDDEQF 64

RESULT 167  
US-09-009-913-339  
; Sequence 339, Application US/09009913  
; Patent No. 6087485  
; GENERAL INFORMATION:

; APPLICANT: AxyS Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Asthma Related Genes  
; NUMBER OF SEQUENCES: 339  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bozicevic & Reed, LLP  
; STREET: 285 Hamilton Ave, Suite 200  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/009,913  
; FILING DATE: 21-JAN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: 36,677  
; REFERENCE/DOCKET NUMBER: SEQ-4P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-327-3231  
; TELEFAX: 650-327-3231  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 339:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 283 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-09-009-913-339

Query Match 1.8%; Score 6; DB 3; Length 283;  
Best Local Similarity 100.0%; Pred. No. 8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69  
Db 243 EAVAQL 248

## RESULT 168

US-09-071-035-426  
; Sequence 426, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 426:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 284 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-071-035-426

Query Match 1.8%; Score 6; DB 4; Length 284;  
Best Local Similarity 100.0%; Pred. No. 8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
QY      79 LVTLLA 84
      |||||
Db     18 LVTLLA 23

RESULT 169
US-09-292-858B-25
; Sequence 25, Application US/09292858B
; Patent No. 645681
; GENERAL INFORMATION:
; APPLICANT: Dean, Frank
; APPLICANT: O'Donnell, Michael E.
; TITLE OF INVENTION: DNA MOLECULES ENCODING SINGLE STRAND GAP RESPONSE
; TITLE OF INVENTION: PROTEINS INVOLVED IN ACTIVATION OF A DNA REPAIR/CELL
; TITLE OF INVENTION: CYCLE CHECKPOINT PATHWAY
; FILE REFERENCE: 22221/1011
; CURRENT APPLICATION NUMBER: US/09/292,858B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,020
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-09-292-858B-25

Query Match      1.8%; Score 6; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      208 TIFEDY 213
      |||||
Db     57 TIFEDY 62

RESULT 170
US-09-543-681A-6438
; Sequence 6438, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6438
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6438

Query Match      1.8%; Score 6; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      134 EAPQIA 139
      |||||
Db     234 EAPQIA 239

RESULT 171
US-09-134-000C-4481
; Sequence 4481, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4481
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4481

Query Match      1.8%; Score 6; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      79 LVTLLA 84
      |||||
Db     28 LVTLLA 33

RESULT 172
US-09-150-133-13
; Sequence 13, Application US/09150133B
; Patent No. 6060295
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.504
; CURRENT APPLICATION NUMBER: US/09/150,133B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
US-09-150-133-13

Query Match      1.8%; Score 6; DB 3; Length 295;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      300 KLIEFL 305
      |||||
Db     206 KLIEFL 211

RESULT 173
US-09-150-141-13
; Sequence 13, Application US/09150141B
; Patent No. 6071732
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES, AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.495
; CURRENT APPLICATION NUMBER: US/09/150,141B
; CURRENT FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 5.1 (saved in ASCII format)
; SEQ ID NO 13
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: estrogen sulfotransferase
```

## US-09-150-141-13

Query Match 1.8%; Score 6; DB 3; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIIEFL 305  
DB 206 KLIIEFL 211

## RESULT 174

US-09-374-493-13  
; Sequence 13, Application US/09374493  
; Patent No. 6204016  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5820.546  
; CURRENT APPLICATION NUMBER: US/09/374,493  
; CURRENT FILING DATE: 1999-08-13  
; EARLIER APPLICATION NUMBER: 09/150,133  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 60/072,994  
; EARLIER FILING DATE: 1998-01-29  
; EARLIER APPLICATION NUMBER: PCT/US99/16750  
; EARLIER FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-374-493-13

Query Match 1.8%; Score 6; DB 3; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIIEFL 305  
DB 206 KLIIEFL 211

## RESULT 175

US-09-374-824-13  
; Sequence 13, Application US/09374824  
; Patent No. 6207414  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5820.547  
; CURRENT APPLICATION NUMBER: US/09/374,824  
; CURRENT FILING DATE: 1999-08-13  
; EARLIER APPLICATION NUMBER: 09/150,133  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 60/072,994  
; EARLIER FILING DATE: 1998-01-29  
; EARLIER APPLICATION NUMBER: PCT/US99/16750  
; EARLIER FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-374-824-13

Query Match 1.8%; Score 6; DB 3; Length 295;

Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIIEFL 305  
DB 206 KLIIEFL 211

## RESULT 176

US-09-374-492-13  
; Sequence 13, Application US/09374492  
; Patent No. 6207432  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5820.545  
; CURRENT APPLICATION NUMBER: US/09/374,492  
; CURRENT FILING DATE: 1999-08-13  
; EARLIER APPLICATION NUMBER: 09/150,141  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 60/072,994  
; EARLIER FILING DATE: 1998-01-29  
; EARLIER APPLICATION NUMBER: PCT/US99/16750  
; EARLIER FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-374-492-13

Query Match 1.8%; Score 6; DB 3; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 KLIIEFL 305  
DB 206 KLIIEFL 211

## RESULT 177

US-09-785-343-13  
; Sequence 13, Application US/09785343  
; Patent No. 6605455  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Regents of the University of Oklahoma  
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5827.003  
; CURRENT APPLICATION NUMBER: US/09/785,343  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: 09/150,133  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/072,994  
; PRIOR FILING DATE: 1998-01-29  
; PRIOR APPLICATION NUMBER: PCT/US99/16750  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)  
; SEQ ID NO 13  
; LENGTH: 295  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: estrogen sulfotransferase  
US-09-785-343-13

Query Match 1.8%; Score 6; DB 4; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 300 KLIFFL 305  
| | | | |  
Db 206 KLIFFL 211

RESULT 178  
US-09-134-000C-3737  
; Sequence 3737, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3737  
; LENGTH: 295  
; TYPE: PR7  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-3737

Query Match 1.8%; Score 6; DB 4; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 TLIADL 86  
| | | | |  
Db 118 TLIADL 123

RESULT 179  
US-08-923-856-1  
; Sequence 1, Application US/08923856  
; Patent No. 5928894  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Tang, Tom  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: HUMAN ACTVA-ORF4-LIKE PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,856  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0380 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-845-4166  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 299 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT01  
; CLONE: 223909  
US-08-923-856-1

Query Match 1.8%; Score 6; DB 2; Length 299;  
Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 EDYEKL 216  
| | | | |  
Db 244 EDYEKL 249

RESULT 180  
US-09-216-294-1  
; Sequence 1, Application US/09216294  
; Patent No. 6080723  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Tang, Tom  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: HUMAN ACTVA-ORF4-LIKE PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/216,294  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/923,856  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0380 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 299 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT01  
; CLONE: 223909  
US-09-216-294-1

Query Match 1.8%; Score 6; DB 3; Length 299;  
Best Local Similarity 100.0%; Pred. No. 8.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 EDYEKL 216  
| | | | |  
Db 244 EDYEKL 249

```
RESULT 181
US-09-107-532A-4050
; Sequence 4050, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4050:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...299
; SEQUENCE DESCRIPTION: SEQ ID NO: 4050:
US-09-107-532A-4050

Query Match 1.8%; Score 6; DB 4; Length 299;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 KVLVAD 200
Db 68 KVLVAD 73

RESULT 182
US-09-009-913-5
; Sequence 5, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axis Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto

Query Match 1.8%; Score 6; DB 4; Length 299;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 KVLVAD 200
Db 68 KVLVAD 73

RESULT 183
US-09-976-594-7
; Sequence 7, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3170236CD1
US-09-976-594-7

Query Match 1.8%; Score 6; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 268 NIOFEA 273
Db 286 NIOFEA 291
```

```
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-009-913-5

Query Match 1.8%; Score 6; DB 3; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 EAVAQL 69
Db 243 EAVAQL 248

RESULT 183
US-09-976-594-7
; Sequence 7, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3170236CD1
US-09-976-594-7

Query Match 1.8%; Score 6; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 268 NIOFEA 273
Db 286 NIOFEA 291
```

```

RESULT 184
US-09-252-991A-27266
; Sequence 27266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27266
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27266

Query Match          1.8%; Score 6; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LVLTLI 83
Db 16 LVLTLI 21

RESULT 185
US-09-328-352-4994
; Sequence 4994, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4994
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4994

Query Match          1.8%; Score 6; DB 4; Length 307;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 YEKLLQ 218
Db 263 YEKLLQ 268

RESULT 186
US-09-461-325-437
; Sequence 437, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16

```

```

; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 437
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-325-437

Query Match          1.8%; Score 6; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 GLLVTL 82
Db 226 GLLVTL 231

RESULT 187
US-09-252-991A-27921
; Sequence 27921, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27921
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27921

Query Match          1.8%; Score 6; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 LRRQIG 110
Db 162 LRRQIG 167

RESULT 188
US-09-252-991A-29243
; Sequence 29243, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29243  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29243

Query Match 1.8%; Score 6; DB 4; Length 308;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SSGLLV 80  
Db 22 SSGLLV 27

## RESULT 189

US-09-328-352-7093  
; Sequence 7093, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7093  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7093

Query Match 1.8%; Score 6; DB 4; Length 308;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 SLKLLG 233  
Db 138 SLKLLG 143

## RESULT 190

US-09-328-352-7158  
; Sequence 7158, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7158  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7158

Query Match 1.8%; Score 6; DB 4; Length 308;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 LYSSGL 78  
Db 150 LYSSGL 155

## RESULT 191

US-10-012-542-437  
; Sequence 437, Application US/10012542

; Patent No. 6627741  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1  
; CURRENT APPLICATION NUMBER: US/10/012,542  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 437  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-012-542-437

Query Match 1.8%; Score 6; DB 4; Length 308;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GLLVTL 82  
Db 226 GLLVTL 231

## RESULT 192

US-09-655-908-12  
; Sequence 12, Application US/09655908  
; Patent No. 6645747  
; GENERAL INFORMATION:  
; APPLICANT: Hallahan, David L.  
; TITLE OF INVENTION: cis-prenyltransferases from Plants  
; FILE REFERENCE: BC1019 US NA  
; CURRENT APPLICATION NUMBER: US/09/655,908  
; CURRENT FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/155,046  
; PRIOR FILING DATE: 1999-09-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Vitis sp  
US-09-655-908-12

Query Match 1.8%; Score 6; DB 4; Length 309;  
Best Local Similarity 100.0%; Pred. No. 8.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 LSSFOK 310  
Db 295 LSSFOK 300

## RESULT 193

US-09-252-991A-22411  
; Sequence 22411, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22411  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22411

Query Match 1.8%; Score 6; DB 4; Length 313;  
Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 83 IADLQL 88  
Db 176 IADLQL 181

RESULT 194  
US-09-489-039A-8342  
; Sequence 8342, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8342  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8342

Query Match 1.8%; Score 6; DB 4; Length 314;  
Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 63 TEVAQ 68  
Db 283 TEVAQ 288

RESULT 195  
US-09-634-137-30  
; Sequence 30, Application US/09634137  
; Patent No. 6632665  
; GENERAL INFORMATION:  
; APPLICANT: Perrino, Fred W  
; TITLE OF INVENTION: Mammalian Genes Encoding 3'-5' Exonuclease  
; FILE REFERENCE: wak200/48001/4-018  
; CURRENT APPLICATION NUMBER: US/09/634,137  
; CURRENT FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: US 60/148,018  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-634-137-30

Query Match 1.8%; Score 6; DB 4; Length 314;  
Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 23 DNLAIL 28  
Db 102 DNLAIL 107

RESULT 196  
US-08-571-758-12  
; Sequence 12, Application US/08571758  
; Patent No. 5700675  
; GENERAL INFORMATION:  
; APPLICANT: Rubin, Gerry M.  
; APPLICANT: Therrien, Marc  
; APPLICANT: Chang, Henry C.  
; APPLICANT: Karim, Felix D.  
; APPLICANT: Wassatman, David A.  
; TITLE OF INVENTION: A No. 5700675el Protein Kinase Required for Ras  
; TITLE OF INVENTION: Signal Transduction  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/571,758  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B96-010  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 315 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-571-758-12

Query Match 1.8%; Score 6; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 22 KDLAI 27  
Db 81 KDLAI 86

RESULT 197  
US-08-909-984A-12  
; Sequence 12, Application US/08909984A  
; Patent No. 5747275  
; GENERAL INFORMATION:  
; APPLICANT: Rubin, Gerry M.  
; APPLICANT: Therrien, Marc  
; APPLICANT: Chang, Henry C.  
; APPLICANT: Karim, Felix D.

; APPLICANT: Wasserman, David A.  
; TITLE OF INVENTION: A No. 5747275el Protein Kinase Required for Ras  
; TITLE OF INVENTION: Signal Transduction  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/909,984A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B96-010  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 315 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; US-08-909-984A-12

Query Match 1.8%; Score 6; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KNLAI 27  
Db 81 KNLAI 86

RESULT 198  
US-08-909-983-12  
; Sequence 12, Application US/08090983  
; Patent No. 5747288  
; GENERAL INFORMATION:  
; APPLICANT: Rubin, Gerry M.  
; APPLICANT: Therrien, Marc  
; APPLICANT: Chang, Henry C.  
; APPLICANT: Karim, Felix D.  
; APPLICANT: Wasserman, David A.  
; TITLE OF INVENTION: A No. 5747288el Protein Kinase Required for Ras  
; TITLE OF INVENTION: Signal Transduction  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/909,983  
; FILING DATE: 12-JUN-1997

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/571,758  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B96-010  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 315 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; US-08-909-983-12

Query Match 1.8%; Score 6; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KNLAI 27  
Db 81 KNLAI 86

RESULT 199  
US-09-107-532A-5596  
; Sequence 5596, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5596:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 317 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:



```
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...317
; SEQUENCE DESCRIPTION: SEQ ID NO: 5596:
US-09-107-532A-5596

Query Match          1.8%; Score 6; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      119 ISAHPH 124
Db      41 ISAHPH 46

RESULT 200
US-09-134-001C-4204
; Sequence 4204, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4204
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4204

Query Match          1.8%; Score 6; DB 4; Length 320;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      119 ISAHPH 124
Db      45 ISAHPH 50

RESULT 201
US-09-848-294-4
; Sequence 4, Application US/09848294
; Patent No. 6479640
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640el
; FILE REFERENCE: CSHL90-04FZA
; CURRENT APPLICATION NUMBER: US/09/848,294
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 09/235,251
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/759,536
; PRIOR FILING DATE: 1996-12-04
; PRIOR APPLICATION NUMBER: 08/107,420
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: 07/663,579
; PRIOR FILING DATE: 1991-03-01
; PRIOR APPLICATION NUMBER: 07/494,036
; PRIOR FILING DATE: 1990-03-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 322
```

```
; TYPE: PRT
; ORGANISM: Homosapiens
US-09-848-294-4

Query Match          1.8%; Score 6; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      75 SGLLV 80
Db      207 SGLLV 212

RESULT 202
US-09-252-991A-18724
; Sequence 18724, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18724
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18724

Query Match          1.8%; Score 6; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      195 KVLVAD 200
Db      69 KVLVAD 74

RESULT 203
US-08-044-812A-4
; Sequence 4, Application US/08044812A
; Patent No. 5837521
; GENERAL INFORMATION:
; APPLICANT: Cone, Roger D
; APPLICANT: Roselli-Reh fuss, Linda
; APPLICANT: Mountjoy, Kathleen G
; APPLICANT: Robbins, Linda S
; TITLE OF INVENTION: Mammalian Melanocyte Stimulating Hormone
; TITLE OF INVENTION: Receptors and Uses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/044,812A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

/ NAME: No. 5837521nan, Kevin E  
/ REGISTRATION NUMBER: 35,303  
/ REFERENCE/DOCKET NUMBER: 92,835  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 312-715-1000  
/ TELEFAX: 312-715-1234  
/ TELEX: 910-221-5317  
/ INFORMATION FOR SEQ ID NO: 4:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 323 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-08-044-812A-4

Query Match 1.8%; Score 6; DB 2; Length 323;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KEILCG 55  
Db 311 KEILCG 316

RESULT 204  
US-08-475-637-4  
/ Sequence 4, Application US/08475637  
/ Patent No. 5994087  
/ GENERAL INFORMATION:  
/ APPLICANT: Cone, Roger D  
/ APPLICANT: Roselli-Rehfuess, Linda  
/ APPLICANT: Mountjoy, Kathleen G  
/ APPLICANT: Robbins, Linda S  
/ TITLE OF INVENTION: Mammalian Melanocyte Stimulating Hormone  
/ TITLE OF INVENTION: Receptors and Uses  
/ NUMBER OF SEQUENCES: 4  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESS: Allegretti & Witcoff, Ltd.  
/ STREET: 10 South Wacker Drive, Suite 3000  
/ CITY: Chicago  
/ STATE: Illinois  
/ COUNTRY: USA  
/ ZIP: 60606

/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/475,637  
/ FILING DATE: 07-JUN-1995  
/ CLASSIFICATION: 435  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/044,812  
/ FILING DATE: 04-APR-1993  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: No. 5994087nan, Kevin E  
/ REGISTRATION NUMBER: 35,303  
/ REFERENCE/DOCKET NUMBER: 92,835  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 312-715-1000  
/ TELEFAX: 312-715-1234  
/ TELEX: 910-221-5317  
/ INFORMATION FOR SEQ ID NO: 4:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 323 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-08-475-637-4

Query Match 1.8%; Score 6; DB 2; Length 323;

Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KEILCG 55  
Db 311 KEILCG 316

RESULT 205  
US-08-706-281A-12  
/ Sequence 12, Application US/08706281A  
/ Patent No. 6100048  
/ GENERAL INFORMATION:  
/ APPLICANT: Cone, Roger D  
/ APPLICANT: Fan, Wei  
/ APPLICANT: Boston, Bruce A  
/ APPLICANT: Kesterton, Robert A  
/ APPLICANT: Lu, Dongai  
/ APPLICANT: Chen, Wenbiao  
/ TITLE OF INVENTION: Methods and Reagents for Discovering and  
/ TITLE OF INVENTION: Using Mammalian Melanocortin Receptor Agonists and Antagonist  
/ TITLE OF INVENTION: To Modulate Feeding Behavior in Animals  
/ NUMBER OF SEQUENCES: 19  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
/ STREET: 300 South Wacker Drive  
/ CITY: Chicago  
/ STATE: IL  
/ COUNTRY: USA  
/ ZIP: 60606

/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/706,281A  
/ FILING DATE: 04-SEP-1996  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: No. 6100048nan, Kevin E  
/ REGISTRATION NUMBER: 35,303  
/ REFERENCE/DOCKET NUMBER: 96,886  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 312-913-0001  
/ TELEFAX: 312-913-0002  
/ TELEX:

/ INFORMATION FOR SEQ ID NO: 12:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 323 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-08-706-281A-12

Query Match 1.8%; Score 6; DB 3; Length 323;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KEILCG 55  
Db 311 KEILCG 316

RESULT 206  
US-09-191-359-4  
/ Sequence 4, Application US/09191359  
/ Patent No. 6261838  
/ GENERAL INFORMATION:  
/ APPLICANT: Cone, Roger D  
/ APPLICANT: Roselli-Rehfuess, Linda  
/ APPLICANT: Mountjoy, Kathleen G  
/ APPLICANT: Robbins, Linda S

;; TITLE OF INVENTION: Mammalian Melanocyte Stimulating Hormone  
;;  
;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Allegretti & Witcoff, Ltd.  
;; STREET: 10 South Wacker Drive, Suite 3000  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/191,359  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,637  
;; FILING DATE: 07-JUN-1995  
;; APPLICATION NUMBER: US/08/044,812  
;; FILING DATE: 04-APR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6261838nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 92,835  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 910-221-5317  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 323 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
;; US-09-191-359-4

Query Match 1.8%; Score 6; DB 3; Length 323;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KEILCG 55  
Db 311 KEILCG 316

RESULT 207  
US-09-097-231-12  
; Sequence 12, Application US/09097231  
; Patent No. 6278038  
; GENERAL INFORMATION:  
; APPLICANT: Cone, Roger D  
; Chen, Wenbiao  
; Low, Malcolm J  
; TITLE OF INVENTION: Mammalian Melanocortin Receptor and Uses  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/097,231

;; FILING DATE: 12-Jun-1998  
;; CLASSIFICATION: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6278038nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 96,886-C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-913-0001  
;; TELEFAX: 312-913-0002  
;; TELEX: <Unknown>  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 323 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
;; US-09-097-231-12

Query Match 1.8%; Score 6; DB 3; Length 323;  
Best Local Similarity 100.0%; Pred. No. 9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 KEILCG 55  
Db 311 KEILCG 316

RESULT 208  
US-09-353-099-12  
; Sequence 12, Application US/09353099  
; Patent No. 6476187  
; GENERAL INFORMATION:  
; APPLICANT: Cone, Roger D  
; Fan, Wei  
; Boston, Bruce A  
; Kesterton, Robert A  
; Lu, Dongsi  
; Chen, Wenbiao  
; TITLE OF INVENTION: Methods and Reagents for Discovering and  
; Using Mammalian Melanocortin Receptor Agonists and Antagon  
; To Modulate Feeding Behavior in Animals

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/353,099  
FILING DATE: 14-Sep-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/706,281  
FILING DATE: 04-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: No. 6476187nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 96,886  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 323 amino acids



; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10843  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-10843

Query Match 1.8%; Score 6; DB 4; Length 331;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 HEPLAK 157  
Db 273 HEPLAK 278

RESULT 214  
US-09-331-581-23  
; Sequence 23, Application US/09331581  
; Patent No. 6130070  
; GENERAL INFORMATION:  
; APPLICANT: TOHDA, Hideki  
; APPLICANT: HANA, Yuko  
; APPLICANT: KUNAGAI, Hiromichi  
; TITLE OF INVENTION: INDUCTION PROMOTER GENE AND SECRETORY SIGNAL GENE  
; TITLE OF INVENTION: USABLE IN SCHIZOSACCHAROMYCES POMBE, EXPRESSION VECTORS  
; FILE REFERENCE: 0059-1142-OPCT  
; CURRENT APPLICATION NUMBER: US/09/331,581  
; CURRENT FILING DATE: 1999-06-30  
; EARLIER APPLICATION NUMBER: PCT/JP98/04929  
; EARLIER FILING DATE: 1998-10-30  
; EARLIER APPLICATION NUMBER: JP 9-314608  
; EARLIER FILING DATE: 1997-10-31  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Schwannomyces occidentalis  
US-09-331-581-23

Query Match 1.8%; Score 6; DB 3; Length 332;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 SNQFRD 167  
Db 164 SNQFRD 169

RESULT 215  
US-09-489-039A-9330  
; Sequence 9330, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9330

; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9330

Query Match 1.8%; Score 6; DB 4; Length 332;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 IVEILL 295  
Db 16 IVEILL 21

RESULT 216  
US-09-134-000C-5207  
; Sequence 5207, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 5207  
; LENGTH: 333  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5207

Query Match 1.8%; Score 6; DB 4; Length 333;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 ISAHPH 124  
Db 50 ISAHPH 55

RESULT 217  
US-09-543-681A-5285  
; Sequence 5285, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 5285  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-5285

Query Match 1.8%; Score 6; DB 4; Length 334;  
Best Local Similarity 100.0%; Pred. No. 9.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 GELILD 238  
Db 38 GELILD 43

```
RESULT 218
US-09-674-826B-6
; Sequence 6, Application US/09674826B
; Patent No. 6638735
; GENERAL INFORMATION:
; APPLICANT: Doosan Corporation
; TITLE OF INVENTION: Korea Institute of Science and Technology
; TITLE OF INVENTION: Plasmid for gene expression in Pichia ciferri and
; FILE REFERENCE: PCT-981031 transformation method using the same
; CURRENT APPLICATION NUMBER: US/09/674,826B
; CURRENT FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Pichia ciferrii
US-09-674-826B-6

Query Match      1.8%; Score 6; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 PLAKII 159
Db 158 PLAKII 163

RESULT 219
US-09-107-532A-6417
; Sequence 6417, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6417:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
```

```
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...336
; SEQUENCE DESCRIPTION: SEQ ID NO: 6417:
US-09-107-532A-6417

Query Match      1.8%; Score 6; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMNLL 262
Db 18 LMNLL 23

RESULT 220
US-09-057-762-2
; Sequence 2, Application US/09057762
; Patent No. 5879909
; GENERAL INFORMATION:
; APPLICANT: PERL, ANDRAS
; TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH
; TITLE OF INVENTION: A FUNCTION IN METABOLISM
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,762
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/326,119
; FILING DATE: 13-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LIVNAT, SHMUEL
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: 280932000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-057-762-2

Query Match      1.8%; Score 6; DB 2; Length 337;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLIGEL 235
Db 258 KLIGEL 263

RESULT 221
US-08-326-119A-2
; Sequence 2, Application US/08326119A
; Patent No. 6018021
```

GENERAL INFORMATION:  
; APPLICANT: PERL, ANDRAS  
; TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH  
; TITLE OF INVENTION: A FUNCTION IN METABOLISM  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/326,119A  
; FILING DATE: 19-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LIVNAT, SHMUEL  
; REGISTRATION NUMBER: 33,949  
; REFERENCE/DOCKET NUMBER: 280932000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 822-0168  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 337 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-326-119A-2

Query Match 1.8%; Score 6; DB 3; Length 337;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 KLLGEL 235  
Db 258 KLLGEL 263

RESULT 222  
US-09-198-452A-642  
; Sequence 642, Application US/09198452A  
; Patent No. 659294  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 642  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-642

Query Match 1.8%; Score 6; DB 4; Length 338;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 324 YLIKQI 329  
Db 28 YLIKQI 33

RESULT 223  
US-09-614-912-150  
; Sequence 150, Application US/09614912  
; Patent No. 6677502  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Rafalski, Antoni  
; APPLICANT: Orozco, Buddy  
; APPLICANT: Miao, Gou-Hau  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian Ming  
; APPLICANT: Sakai, Hajime  
; APPLICANT: Weng, Zude  
; APPLICANT: Calmi, Perry G  
; APPLICANT: Anderson, Shawn  
; TITLE OF INVENTION: Plant Metabolism Genes  
; FILE REFERENCE: BB1378 US NA  
; CURRENT APPLICATION NUMBER: US/09/614,912  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: 60/143,401  
; PRIOR FILING DATE: 1999-07-12  
; PRIOR APPLICATION NUMBER: 60/143,412  
; PRIOR FILING DATE: 1999-07-12  
; PRIOR APPLICATION NUMBER: 60/146,650  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 60/170,906  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: 60/172,959  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/172,946  
; PRIOR FILING DATE: 1999-12-21  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 150  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-614-912-150

Query Match 1.8%; Score 6; DB 4; Length 338;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 DNLAIL 28  
Db 87 DNLAIL 92

RESULT 224  
US-08-276-151-5  
; Sequence 5, Application US/08276151  
; Patent No. 5597719  
; GENERAL INFORMATION:  
; APPLICANT: Freed, Ellen  
; APPLICANT: Ruggieri, Rosamaria  
; TITLE OF INVENTION: Interaction of raf-1 and 14-3-3 Proteins  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooley Godward et al.  
; STREET: Five Palo Alto Square  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/276,151  
; FILING DATE: 14-JUL-1994

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: ONYX-005/00US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 843-5481  
TELEFAX: (415) 857-0663  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: C-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-276-151-5

Query Match 1.8%; Score 6; DB 1; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KNLAI 27  
DB 112 KNLAI 117

RESULT 225  
US-08-476-254-2  
Sequence 2, Application US/08476254  
Patent No. 5846531  
GENERAL INFORMATION:  
APPLICANT: WEINER, RONALD M.  
APPLICANT: FUQUA, WILLIAM C.  
TITLE OF INVENTION: MARINE MELA GENE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WATSON COLE STEVENS DAVIS, P.L.L.C.  
STREET: 1400 K STREET NW  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-2477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,254  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: POULOS III, JAMES A.  
REGISTRATION NUMBER: 31,714  
REFERENCE/DOCKET NUMBER: JAP30319C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 628-0088  
TELEFAX: 202 628-8034  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-254-2

Query Match 1.8%; Score 6; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 YDTIFE 211  
DB 267 YDTIFE 272

RESULT 226  
US-08-476-254-10  
Sequence 10, Application US/08476254  
Patent No. 5846531  
GENERAL INFORMATION:  
APPLICANT: WEINER, RONALD M.  
APPLICANT: FUQUA, WILLIAM C.  
TITLE OF INVENTION: MARINE MELA GENE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WATSON COLE STEVENS DAVIS, P.L.L.C.  
STREET: 1400 K STREET NW  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-2477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,254  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: POULOS III, JAMES A.  
REGISTRATION NUMBER: 31,714  
REFERENCE/DOCKET NUMBER: JAP30319C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 628-0088  
TELEFAX: 202 628-8034  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: S. colwelliana  
STRAIN: D  
US-08-476-254-10

Query Match 1.8%; Score 6; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 YDTIFE 211  
DB 267 YDTIFE 272

RESULT 227  
US-09-153-599A-10  
Sequence 10, Application US/09153599A  
Patent No. 6420177  
GENERAL INFORMATION:  
APPLICANT: Weber, J. Mark  
APPLICANT: Luu, B. Minh  
TITLE OF INVENTION: Method for Strain Improvement of  
TITLE OF INVENTION: Erythromycin Producing Bacterium  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rocky, Milnamow & Katz, Ltd.  
STREET: 180 N. Stetson Avenue, 2 Prudential Plaza



; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/153,599A  
; APPLICATION NUMBER: US/09/153,599A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V.  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: PER159P0041US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 346 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-153-599A-10

Query Match 1.8%; Score 6; DB 4; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 YDTIFE 211  
Db 267 YDTIFE 272

RESULT 228  
5474933-2  
; Patent No. 5474933  
; APPLICANT: EINER, RONALD M.; FUQUA JR., WILLIAM C.  
; TITLE OF INVENTION: MARINE MELA GENE  
; NUMBER OF SEQUENCES: 9  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/148,945  
; FILING DATE: 08-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 974,837  
; FILING DATE: 10-NOV-1992  
; APPLICATION NUMBER: 496,804  
; FILING DATE: 21-MAR-1990  
; SEQ ID NO: 2:  
; LENGTH: 346  
5474933-2

Query Match 1.8%; Score 6; DB 6; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 YDTIFE 211  
Db 267 YDTIFE 272

RESULT 229  
5474933-7  
; Patent No. 5474933  
; APPLICANT: EINER, RONALD M.; FUQUA JR., WILLIAM C.  
; TITLE OF INVENTION: MARINE MELA GENE  
; NUMBER OF SEQUENCES: 9  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/148,945

; FILING DATE: 08-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 974,837  
; FILING DATE: 10-NOV-1992  
; APPLICATION NUMBER: 496,804  
; FILING DATE: 21-MAR-1990  
; SEQ ID NO: 7:  
; LENGTH: 346  
5474933-7

Query Match 1.8%; Score 6; DB 6; Length 346;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 YDTIFE 211  
Db 267 YDTIFE 272

RESULT 230  
US-09-252-991A-31854  
; Sequence 31854, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31854  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31854

Query Match 1.8%; Score 6; DB 4; Length 347;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 POIALR 141  
Db 141 POIALR 146

RESULT 231  
US-09-489-039A-7292  
; Sequence 7292, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 7292  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7292

Query Match 1.8%; Score 6; DB 4; Length 351;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      82 LIADLQ 87
Db      343 LIADLQ 348

RESULT 232
US-09-198-452A-702
; Sequence 702, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 702
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-702

Query Match      1.8%; Score 6; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      81 TLIADL 86
Db      58 TLIADL 63

RESULT 233
US-09-328-352-5230
; Sequence 5230, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GT99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5230
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5230

Query Match      1.8%; Score 6; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      214 EKLIQS 219
Db      168 EKLIQS 173

RESULT 234
US-09-489-039A-9887
; Sequence 9887, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
```

```
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9887
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9887

Query Match      1.8%; Score 6; DB 4; Length 355;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      231 LIGELI 236
Db      282 LIGELI 287

RESULT 235
US-09-134-000C-4914
; Sequence 4914, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4914
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4914

Query Match      1.8%; Score 6; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      77 GLLVTL 82
Db      277 GLLVTL 282

RESULT 236
US-08-978-182-4
; Sequence 4, Application US/08978182
; Patent No. 5849556
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Kaser, Matthew
; APPLICANT: Mathur, Preete
; TITLE OF INVENTION: HUMAN GROWTH-RELATED CDC10 HOMOLOG
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,182
```

```
/ FILING DATE: Herewith
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0426 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 357 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 578115
/ US-08-978-182-4

Query Match 1.8%; Score 6; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 KLLQSE 220
Db 198 KLLQSE 203

RESULT 237
US-09-205-681-4
/ Sequence 4, Application US/09205681
/ Patent No. 5952214
/ GENERAL INFORMATION:
/ APPLICANT: Hillman, Jennifer L.
/ APPLICANT: Yue, Henry
/ APPLICANT: Guegler, Karl J.
/ APPLICANT: Kaser, Matthew
/ APPLICANT: Mathur, Preetee
/ TITLE OF INVENTION: HUMAN GROWTH-RELATED CDC10 HOMOLOG
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/205,681
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/978,182
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0426 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
```

```
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 357 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 578115
/ US-09-205-681-4
```

Query Match 1.8%; Score 6; DB 2; Length 357;  
Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 KLLQSE 220  
Db 198 KLLQSE 203

```
RESULT 238
US-09-634-955B-21
/ Sequence 21, Application US/09634955B
/ Patent No. 6511834
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel
/ APPLICANT: Cook, William James
/ TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE
/ TITLE OF INVENTION: MOLECULES AND USES THEREFOR
/ FILE REFERENCE: MNI-134
/ CURRENT APPLICATION NUMBER: US/09/634,955B
/ PRIOR FILING DATE: 2000-08-08
/ PRIOR FILING DATE: 2000-03-24
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: FastSeq for Windows Version 4.
/ SEQ ID NO 21
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: NAD-dependent epimerase/dehydratase domain
/ US-09-634-955B-21
```

Query Match 1.8%; Score 6; DB 4; Length 359;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 KLLGEL 235  
Db 292 KLLGEL 297

```
RESULT 239
US-09-252-991A-21741
/ Sequence 21741, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 21741
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-21741
```

Query Match 1.8%; Score 6; DB 4; Length 359;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 LVADFL 202  
| | | | |  
DB 272 LVADFL 277

RESULT 240  
US-09-080-897-6  
; Sequence 6, Application US/09080897  
; Patent No. 5985574  
; GENERAL INFORMATION:  
; APPLICANT: King, Mary-Claire  
; APPLICANT: Lynch, Eric D.  
; APPLICANT: Lee, Ming  
; APPLICANT: Morrow, Jan E.  
; APPLICANT: Welcsh, Piri L.  
; APPLICANT: Leon, Pedro E.  
; TITLE OF INVENTION: Modulators of Actin  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 75 DENISE DRIVE  
; CITY: HILLSBOROUGH  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94010  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/080,897  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UW97-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 343-4341  
; TELEFAX: (650) 343-4342  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-080-897-6

Query Match 1.8%; Score 6; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 BIVKIL 21  
| | | | |  
DB 153 BIVKIL 158

RESULT 241  
US-09-323-735-6  
; Sequence 6, Application US/09323735  
; Patent No. 6197932  
; GENERAL INFORMATION:  
; APPLICANT: King, Mary-Claire  
; APPLICANT: Lynch, Eric D.  
; APPLICANT: Lee, Ming  
; APPLICANT: Morrow, Jan E.

; APPLICANT: Welcsh, Piri L.  
; APPLICANT: Leon, Pedro E.  
; TITLE OF INVENTION: Modulators of Actin  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 75 DENISE DRIVE  
; CITY: HILLSBOROUGH  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94010  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/323,735  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/080,897  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UW97-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 343-4341  
; TELEFAX: (650) 343-4342  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-323-735-6

Query Match 1.8%; Score 6; DB 3; Length 362;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 BIVKIL 21  
| | | | |  
DB 153 BIVKIL 158

RESULT 242  
US-09-594-669-6  
; Sequence 6, Application US/09594669  
; Patent No. 6331424  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042  
; CURRENT APPLICATION NUMBER: US/09/594,669  
; CURRENT FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: US 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Human  
US-09-594-669-6

Query Match 1.8%; Score 6; DB 4; Length 362;

Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 LKKNQP 299  
Db 128 LKKNQP 133

RESULT 243  
US-10-112-432-6  
; Sequence 6, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6638754el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/112,432  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 09/724,215  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Human

US-10-112-432-6  
Query Match 1.8%; Score 6; DB 4; Length 362;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 294 LKKNQP 299  
Db 128 LKKNQP 133

RESULT 244  
US-09-252-991A-27806  
; Sequence 27806, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27806  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27806

Query Match 1.8%; Score 6; DB 4; Length 363;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 316 EQFADE 321

Db 172 EQFADE 177

RESULT 245  
US-09-489-039A-13446  
; Sequence 13446, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 13446  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-13446

Query Match 1.8%; Score 6; DB 4; Length 363;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 25 LAILEX 30  
Db 22 LAILEX 27

RESULT 246

US-09-107-532A-5920  
; Sequence 5920, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A. Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5920:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 370 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (8) LOCATION 1...370  
SEQUENCE DESCRIPTION: SEQ ID NO: 5920:  
US-09-107-532A-5920

Query Match 1.8%; Score 6; DB 4; Length 370;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 IMTKYI 249  
DB 145 IMTKYI 150

## RESULT 247

US-08-746-789A-2  
Sequence 2, Application US/08746789A  
Patent No. 5789200  
GENERAL INFORMATION:  
APPLICANT: Ismail Kola, Martin J. Tymms, Christine DeBouck  
TITLE OF INVENTION: A No. 5789200el Human ETS Family Member, ELF3  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road, P.O. Box 1539  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
COMPUTER: IBM 486  
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
SOFTWARE: MICROSOFT WORD  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/746.789A  
FILING DATE: No. 5789200ember 15, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: William T. Han  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: ATG 50024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610 270 5219  
TELEFAX: 610 270 4026  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 371  
TYPE: Amino Acid  
TOPOLOGY: Linear

## US-08-746-789A-2

Query Match 1.8%; Score 6; DB 1; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69  
DB 309 EAVAQL 314

## RESULT 248

US-09-570-593-5  
Sequence 5, Application US/09570593  
Patent No. 6566063

GENERAL INFORMATION:  
APPLICANT: Kaufmann, Joerg  
APPLICANT: Xin, Hong  
APPLICANT: Harrowe, Greg  
TITLE OF INVENTION: EXPRESSION OF ETS-DOMAIN PROTEINS IN  
FILE REFERENCE: 2300-1556  
CURRENT APPLICATION NUMBER: US/09/570.593  
CURRENT FILING DATE: 2000-05-12  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Fast-SEQ for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-570-593-5

Query Match 1.8%; Score 6; DB 4; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 EAVAQL 69  
DB 309 EAVAQL 314

## RESULT 249

US-09-543-681A-6429  
Sequence 6429, Application US/09543681A  
Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETTON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI  
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543.681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 6429  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-6429

Query Match 1.8%; Score 6; DB 4; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 EAFHVF 277  
DB 200 EAFHVF 205

## RESULT 250

US-09-489-039A-8389  
Sequence 8389, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489.039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 8389  
LENGTH: 373

```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8389
```

```
Query Match      1.8%; Score 6; DB 4; Length 373;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      22 KDNLA1 27
Db      210 KDNLA1 215
```

```
Search completed: April 12, 2004, 10:37:51
Job time : 29 secs
```